PHD

Strategies for producing therapeutic antibodies in yeast

Adams, Claire Lynnette

Award date:
1999

Awarding institution:
University of Bath

STRATEGIES FOR PRODUCING THERAPEUTIC ANTIBODIES IN YEAST

submitted by Claire Lynnette Adams
for the degree of PhD
of the University of Bath
1999

C. Adams

20.7.99.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Don't Quit

When things go wrong as they sometimes will,
When the road you're trudging seems all up hill.
When the funds are low and the debts are high
And you want to smile, but you have to sigh,
When care is pressing you down a bit,
Rest, if you must but don't you quit.
Life is queer with its twists and turns,
As everyone of us sometimes learns,
And many a failure turns about
When he might have won had he stuck it out;
Don't give up though the pace seems slow-
You may succeed with another blow.
Success is failure turned inside out-
The silver tint of the clouds of doubt,
And you never can tell how close you are,
It may be near when it seems so far;
So stick to the fight when you're hardest hit-
It's when things seem worst that you must not quit.

Anon.

ABSTRACT

The overall aim of this project is to produce high affinity human antibodies that can bind tightly and rapidly to their therapeutic targets. There are many problems associated with producing high affinity monoclonal human antibodies for therapy using established methods such as hybridoma technology, phage display and humanization, therefore another technique is desirable. It was intended to generate high affinity antibodies using a novel *Saccharomyces cerevisiae* antibody improvement strategy that was based upon the *in vivo* germinal centre reaction. Within the germinal centre, antibody improvement (Affinity Maturation) occurs through rounds of random mutation and antigen-selection. Hence, the overall intention of this project was to copy these processes within a yeast cell.

Firstly, to achieve this aim, antibody fragments were expressed in an intact form on the surface of *S. cerevisiae*. Surface expression was achieved by attaching the DNA of an anti-fluorescein-5-isothiocyanate (FITC) single chain Fv antibody fragment to the cell wall anchorage domain of the *S. cerevisiae* α-agglutinin gene. Functional expression of the anti-FITC single chain Fv antibody fragment was detected directly using epi-fluorescence microscopy.

The second part of this project describes the construction of an original antigen-selection strategy, which utilizes the mating characteristics of agglutinin mutants. To overcome variable surface expression, the relevant antibody or antigen DNA was integrated into the appropriate agglutinin loci using a novel Cre-*lox* site-specific recombination system. The engineered α- and **a**-strains were individually analysed for functional expression of the α-agglutinin/antibody and **a**-agglutinin/antigen protein fusions, respectively. Finally, these engineered haploid cells were mated under conditions that were dependent on the antigen-antibody interaction for diploid formation. Unfortunately, this yeast antigen-selection system was unable to differentiate between α-cells expressing the appropriate antibody fragment and non-expressors.

The experiments described in this project will form the basis of future work aimed at establishing an 'in-cell' affinity maturation system in yeast.

ACKNOWLEDGEMENTS

Firstly, I would like to thank my supervisor, Jonathan Cox, for his constant advice and encouragement throughout this project. In addition, I would like to thank Greg Winter, Robert Hawkins, Stephen Oliver, Hans de Nobel and David Jeenes for providing the plasmid DNA described in the relevant sections of this project. The yeast strains YMTAB, FY10 and FY22, were generous gifts from D. Wolff and Fred Winston, respectively. Thanks also to Alan Wheals and Habib Alloush for their general yeast expertise and for plasmid, pFA6a-*kanMX4.* Furthermore, I would like to thank Bill Bennett and Chris Davey for their technical assistance with the fluorescence microscope and Stephen Searle for his help with the structural location of the mutation discovered in antibody fragment, single chain Fv D1.3. To all members of Lab 0.34 (past and present) thank you for making the last three years light-hearted and enjoyable. Last but not least, I would like to express my gratitude to my family and friends, especially my mum and boyfriend Derek, for their continual support and affection.

I would like to thank the Engineering and Physical Sciences Research Council (EPSRC) for funding this work.

TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION

CHAPTER 3 CONSTRUCTION OF A YEAST SURFACE DISPLAY SYSTEM

ABBREVIATIONS

Amino acids

Alanine	Ala	A	Leucine	Leu	L
Arginine	Arg	R	Lysine	Lys	K
Asparagine	Asn	N	Methionine	Met	M
Aspartic Acid	Asp	D	Phenylalanine	Phe	F
Cysteine	Cys	C	Proline	Pro	P
Glutamine	Gln	Q	Serine	Ser	S
Glutamic Acid	Glu	E	Threonine	Thr	T
Glycine	Gly	G	Tryptophan	Trp	W
Histidine	His	H	Tyrosine	Tyr	Y
Isoleucine	Ile	I	Valine	Val	V

Chemicals

ATP	adenosine 5'-triphosphate
BSA	bovine serum albumin
DTT	dithiothreitol
EDTA	disoduim ethylenediamine tetraacetate
FITC	fluorescein-5-isothiocyanate
G418	geneticin
IPTG	isopropyl-β-D-thiogalactosidase
ONPG	o-nitrophenyl-β-D-galactopyranoside
PEG	polyethylene glycol
PMSF	phenylmethylsulphonyl fluoride
pNPP	p-nitrophenyl phosphate
SDS	sodium dodecyl sulphate
TEMED	N, N, N', N'-tetramethylethylenediamine
TMB	tetramethylbenzidine
Tris	tris-(hydroxymethyl)-methylamine
Tween 20	polyoxyethylenesorbitan monolaurate

General

Å	ångstrom unit, 1Å = 0.1nm
Ab	antibody
Abs_{280}	absorbance at 280nm
Ag	antigen
AGA1	**a**-agglutinin anchorage domain locus
AGA2	**a**-agglutinin binding domain locus
APC	antigen-presenting cell
ARS	autonomously replicating sequence
b.p.	base pairs
C	constant
cDNA	complementary DNA
CDR	complementarity-determining region
CEN	centromere
Crick	DNA strand running from the right to the left telomere
D	diversity
dATP	deoxyadenosine 5'-triphosphate
dCTP	deoxycytidine 5'-triphosphate
dGTP	deoxyguanosine 5'-triphosphate
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside 5'-triphosphate
dTTP	deoxythymidine 5'-triphosphate
DZ	dark zone
EIA	enzyme immunoassay
ELISA	enzyme-linked immunosorbent assay
Endo H	endoglycosidase H
ER	endoplasmic reticulum
Fab	fragment antigen-binding
FACS	fluorescence-activated cell sorting
Fc	fragment crystallisable
FDC	follicular dendritic cell
FR	framework region
GAL1	galactose-inducible promoter
GPI	glycosylphosphatidylinositol

H	heavy
HEWL	hen egg white lysozyme
HPLC	high pressure liquid chromatography
Iccosomes	immune complex-coated bodies
Ig	immunoglobulin
J	joining
Ka	association constant
k.b.p.	kilobase pairs
Kd	dissociation constant
kDa	kilodaltons
L	light
LB	Luria Bertani
LMP	low melting point
LZ	light zone
Marvel	non-fat dried milk
MCS	multiple cloning site
mRNA	messenger RNA
n.t.	nucleotide
ORF	open reading frame
ORI	origin of replication
OZ	outer zone
PBS	phosphate buffered saline
PCR	polymerase chain reaction
RNA	ribonucleic acid
r.p.m.	revolutions per minute
SAG1	α-agglutinin locus
SC	synthetic complete
ScFv	single chain fragment variable
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
SGD	*Saccharomyces cerevisiae* genome database
SOE	splicing by overlap extension
TAE	40mM Tris-acetate, 1mM EDTA
Taq	thermostable DNA polymerase from *Thermus aquaticus*

TBE	90mM Tris-borate, 2mM EDTA
UV	ultra-violet
V	variable
v/v	ratio of volume to volume
$Vent_R$	thermostable DNA polymerase from *Thermococcus litoralis*
Watson	DNA strand from the left to the right telomere
w/v	ratio of weight to volume
YPD	yeast peptone dextrose
Δ	deletion
::	novel junction (fusion or insertion)

Bases

A	Adenine
C	Cytosine
G	Guanine
T	Thymine

CHAPTER 1

INTRODUCTION

1.1 THE IMMUNE RESPONSE

The environment contains a large variety of infectious microbial agents; viruses, bacteria, fungi and parasites, all of which can cause great pathological damage if they multiply undetected. The human immune system which combats these agents can be divided into two functional divisions; adaptive and innate immunity, but these do not act in isolation (Roitt *et al.*, 1986). Innate immunity acts as the first line of defence against infectious agents and contains physical barriers such as skin, mucus membranes and circulating elements including complement and phagocytic cells.

If the innate system is not sufficient at preventing infection, the adaptive immune system is activated which produces a specific reaction to each infectious agent. A major element of the adaptive immune system is lymphocytes, which can be divided into two predominant classes, B- and T-cells, which carry out the humoral and cellular components, respectively. During their development, both T- and B-lymphocytes acquire specific receptors for antigen that commits them to a single antigenic specificity for the rest of their life span. The cells are activated when they bind to their specific antigen, proliferate and mature into effector cells (this is known as clonal selection).

The membrane-bound receptors on the surface of T-cells recognise antigens that have been endocytosed by antigen-presenting cells (APCs) such as B-cells and macrophages. Recognition triggers T-cell activation, which according to the class of T-cells causes immune responses such as cytokine release, cytotoxic T-cell activation or differentiation and proliferation of B-cells. Potentially auto-reactive T-cells are removed in the

thymus and this therefore contributes to the discrimination between self and non-self antigens in the immune system (a process known as self-tolerance).

B-cells produce specific antigen receptors known as antibodies that are distinct but structurally related to T-cell receptors. A large repertoire of antibodies is produced by B-cells, which recognise and bind to native antigen. Upon recognition of antigen, B-cells either differentiate into short-lived (plasma) cells which produce soluble antibody or long-lived (memory) cells with membrane-bound antibody. Activation of antibody proliferation triggers a cascade of effects resulting in stimulation of complement and/or recruitment of specialised effector cells including macrophages, phagocytes and natural killer cells. The outcome of these interactions is complex but eventually the overall aim of removing the antigen is achieved.

1.1.1 Physiology of the immune system

The cells involved in the immune system are organized into tissues and organs in order to perform their functions most effectively (Figure 1.1). These structures are collectively referred to as the lymphoid system and comprise of primary and secondary lymphoid organs (Roitt *et al.*, 1986). The primary lymphoid organs are the major sites of lymphocyte differentiation. In humans, T-cells are produced in the thymus and B-cells in the bone marrow and foetal liver. The secondary lymphoid organs (spleen, lymph nodes and tonsils) contain mature T- and B-cells and create an environment in which the lymphocytes can interact with each other and antigen. The lymph nodes form part of the network (lymphatic system) which filters antigen from the tissue fluid during its passage through the body. These comprise of a B-cell area (cortex), a T-cell area (paracortex) and a central medulla. After antigen stimulation, the cortex of the lymph nodes contains aggregates of B-cells known as secondary follicles. The secondary follicles are areas of intense B-lymphocyte proliferation (Section 1.3.3) and can also be found in the periarteriolar lymphoid sheath of the spleen and the mucosal lymphoid tissue of the tonsils.

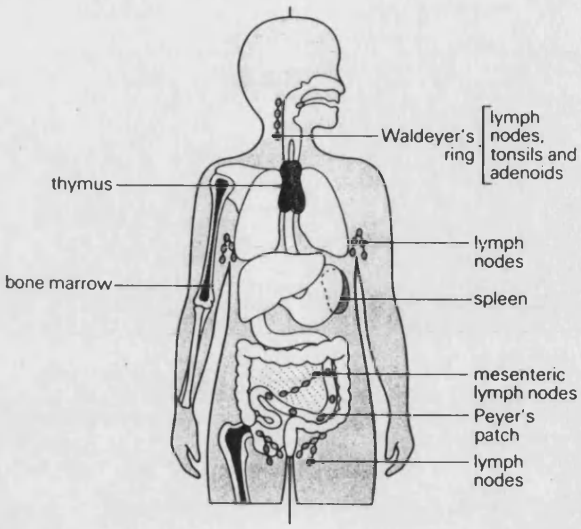


Figure 1.1 - Major lymphoid organs and tissues.

The primary lymphoid tissues, bone marrow and thymus, are represented in the darker grey and produce B- and T-cells, respectively. Lymph nodes are present throughout the human body and only a few are depicted here. Peyer's patches are unencapsulated masses of lymphoid tissue in the small intestine.

1.2 ANTIBODIES

1.2.1 Antibody structure

The structure of an antibody combines two functions: antigen recognition and stimulation of effector function in a single multi-domain molecule (Roitt *et al.*, 1986). In humans, there are five classes of antibody molecules [otherwise known as immunoglobulins (Ig)], namely IgG, IgA, IgM, IgD and IgE. These differ from each other in size, charge, amino acid composition and carbohydrate content, but all have similar overall structures. The class of an immunoglobulin molecule (isotype) is determined by its heavy chain. This project concerns the engineering of the IgG molecule and therefore will concentrate on this major class of antibody.

The IgG molecule consists of two identical light (25kDa) and two identical heavy polypeptide chains (50kDa) linked together by covalent (disulphide) and non-covalent bonds within and between the chains to form an Y-shape (Figure 1.2). This Y-shaped molecule consists of two Fab

(Fragment antigen binding) arms which bind antigen and an Fc (Fragment crystallisable) stalk which interacts with the Fc receptors on different cell types to initiate effector functions such as the complement cascade and general inflammatory responses. Each light chain consists of an N-terminal variable domain (VL) and a C-terminal constant domain (CL). The heavy chain comprises of one variable domain (VH) and three constant domains (CH1, CH2, and CH3). The principal sequence diversity is found in three short stretches within the variable region of the light and heavy chains. These stretches are called hypervariable regions (or complementarity determining regions: CDR1, CDR2, CDR3) and are held in place by four conserved framework regions (FR1, FR2, FR3 and FR4).

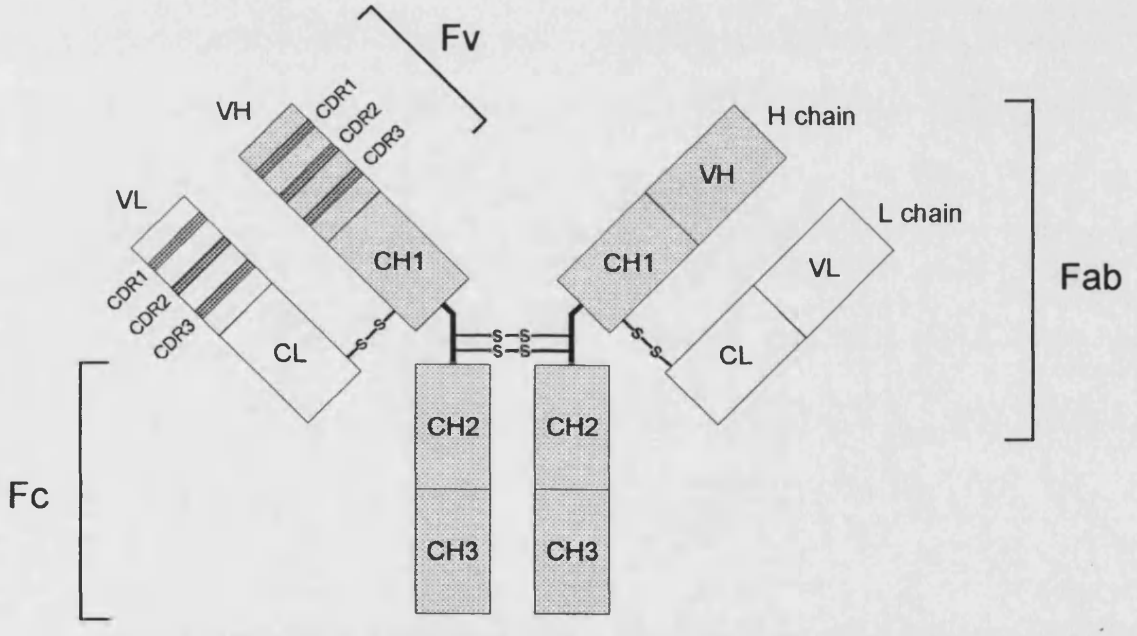


Figure 1.2 - An outline of a typical IgG molecule.

Each IgG molecule is composed of four chains, two heavy (H) and two light (L) chains. The heavy chain comprises of a variable heavy domain (VH) and three constant heavy domains (CH1, CH2 and CH3). The light chain consists of one variable light domain (VL) and one constant light domain (CL). The bivalent Y-shape is formed from specific disulphide bonds between the chains. The Fab region comprises of the Fv fragment (VL+VH) and part of the constant region (CL+CH1), while the Fc region consists of two pairs of constant domains (2 x CH2+CH3).

The three-dimensional (3D) structure of antibodies has been the subject of numerous investigations. The crystal structures for intact immunoglobulins and various antibody fragments are available in both the free state and bound to antigens (Poljak *et al.*, 1973; reviewed in Alzari *et al.*, 1988; Davies *et al.*, 1990). All antibody domains are composed of a characteristic β-sheet sandwich known as an immunoglobulin fold (Figure 1.3). The variable domains consist of nine β-strands and the constant domains contain seven β-strands, all of which are folded in a similar way. The complementarity determining regions form the loops connecting the strands of the variable β-sheets. Non-covalent association of the VL and VH domains produces a β-barrel structure that places the six CDRs in close proximity to form the antigen-combining site. The major determinants in the formation of the antibody tertiary structure are the residues buried within and between the domains (Chothia and Lesk, 1987). Flexibility of the molecule is achieved by the hinge region between the CH1 and CH2 domains (Figure 1.2). This region contains conserved cysteine residues, which are necessary for the formation of disulphide bridges between the various constant domains (Lesk and Chothia, 1982). The number of interchain disulphide bridges vary in different isotypes.

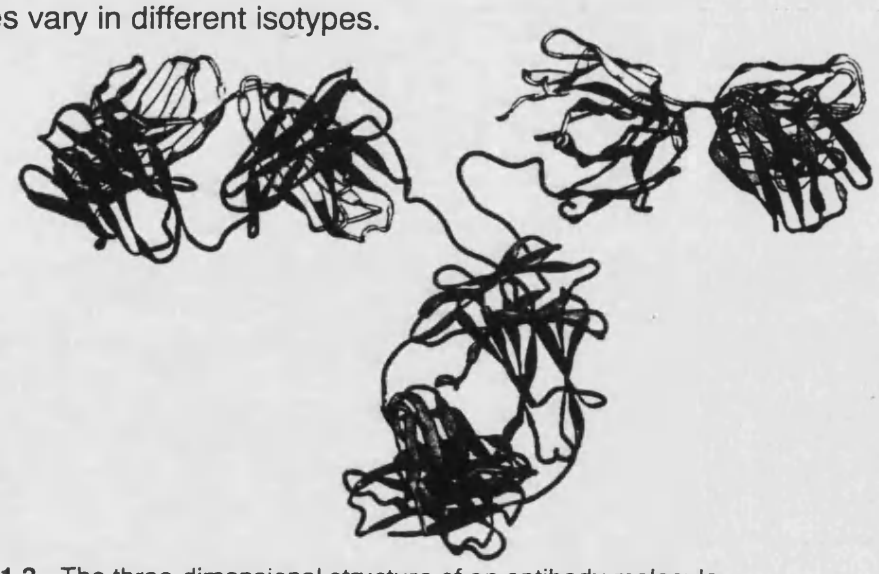


Figure 1.3 - The three-dimensional structure of an antibody molecule.
The heavy chains are shown in grey and the light chains are shown in white. Non-covalent association of the VL and VH domains produces a β-barrel structure that places the CDRs in close proximity to form the antigen-combining site. Figure reproduced from Borrebaeck, 1995.

Generally, the variability and specificity of the antibody is determined by its CDRs and the framework regions conserve the 3D structure. Occasionally, the outer sequences of the framework regions can also affect specificity (Foote and Winter, 1992, Tomlinson *et al.*, 1996). Five of the six complementarity determining regions can be classified into a small structural repertoire of main-chain conformations (otherwise known as canonical structures). These conformations are determined by a few key residues that influence the hydrogen bonding, packing and torsion angles of the canonical structures. Sequence changes at these conserved residues can convert one canonical structure to another (Chothia and Lesk, 1987; Chothia *et al.*, 1989). Almost all of the VH gene segments which contain the first two hypervariable regions exist as one of seven canonical structures (Chothia *et al.*, 1992). Only the conformation of the heavy chain CDR3 can not be predicted by this method due to the variety of size and sequence generated by various genetic mechanisms (Section 1.3.2). Recently, rules relating sequence to conformation of the heavy chain CDR3 have been stated, but to a less complete degree than the other five antigen-binding loops (Morea *et al.*, 1998).

1.2.2 Antibody-antigen interactions

Numerous thermodynamic factors, both enthalpic and entropic affect the association between the antigen and antibody. Considerable configurational (translational and rotational) entropy is lost by each molecule when association occurs. To overcome these barriers van der Waals interactions, hydrogen bonds and salt bridges form between the antibody-antigen interface (Borrebaeck, 1995). These intermolecular attractive forces are optimised by complementary surfaces between the antigenic determinant and the binding site of the antibody. Depressions on one surface are filled by protuberances from the other, often with the total exclusion of water molecules from the interface (Davies *et al.*, 1990). Sometimes, ordered water molecules are found at antigen-antibody interfaces and contribute to the stability and complementarity between the interacting surfaces (Bhat *et*

al, 1994). Local conformational changes by the CDRs of the antibody also improve the efficiency of the association reaction and supports the 'induced fit' hypothesis (Herron *et al.*, 1991; Bhat *et al.*, 1994).

1.3 ANTIBODY IMPROVEMENT *IN VIVO*

To be effective the immune system must be continuously generating diversity to deal with the many potentially harmful infectious and toxic substances in the environment. The primary repertoire in humans comprises of more than 10^9 lymphocytes each displaying a single different antibody (Winter, 1998). This repertoire of antibodies is created by rearrangement of multiple gene segments during B cell development (Section 1.3.2). The affinity of antibodies in this naive population of B cells is normally low (K_a = 10^5-10^7 M^{-1}) but increases following secondary antigenic challenge [(K_a > 10^9 M^{-1}) Winter and Milstein, 1991]. The increase in affinity over the course of an antibody response is due to a process known as affinity maturation (Berek and Milstein, 1987); the rearranged germline variable genes are mutated and antibodies with improved affinity are selected for (Section 1.3.3). Both the recombination and mutation genetic mechanisms greatly diversify the relatively small amount of genetic information encoded by the human genome.

1.3.1 Human immunoglobulin gene organization

The human immunoglobulin chains are encoded by three unlinked gene families; heavy, lambda (λ) light and kappa (κ) light chains which reside on chromosomes 14 (Cook and Tomlinson, 1995), 22 (Frippiat *et al.*, 1995) and 2 (Schäble and Zachau, 1993), respectively. These families are organised in different ways on their particular chromosomes and comprise of distinct gene segments that encode for the variable and constant domains of the antibody molecule.

The Ig heavy gene locus consists of a number of germline gene segments known as variable (V), constant (C), diversity (D) and joining (J).

Each of these segments is separated by non-coding DNA (introns) and are spliced together by a process known as DNA rearrangement (Section 1.3.2) to produce a functional gene. The human heavy chain locus is 1100 kilobases in length and is arranged (5' - 3') with multiple variable exons followed by numerous diversity, joining, and constant gene segments (Figure 1.4).

The human Ig light chain loci, lambda and kappa consist of variable, joining and constant gene segments. The kappa light chain locus is arranged similarly to the heavy chain locus, but each constant region segment is associated with a joining segment in the lambda light chain locus (Figure 1.4).

Conservation at the 5' and 3' ends of the variable germline sequences (i.e. the 5' leader-variable intron region and the 3' recombination signal sequence) has enabled researchers to amplify, clone and sequence the VH, Vκ and Vλ repertoires (Table 1.1). From this data the variable germline gene segments can be divided into different families due to sequence homology. The Vκ and VH gene families are randomly distributed within their relevant loci but the Vλ gene family is organised into three gene clusters (Schäble and Zachau, 1993; Cox *et al.*, 1994; Cook and Tomlinson, 1995; Frippiat *et al.*, 1995).

Table 1.1 - The number of germline gene segments present in the human genome.

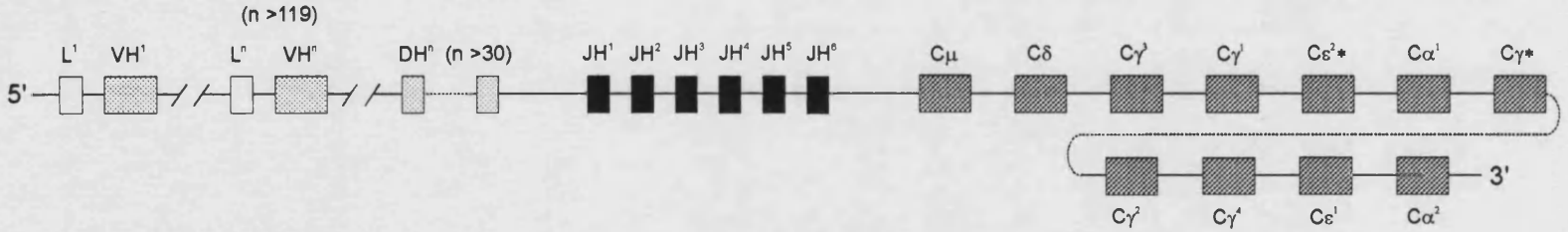
Germline gene segments	H	λ	κ
Variable[a]	>87	>52	>76
Joining	6	7	5
Diversity	>30	0	0
Constant	9	7[b]	1

a - These figures are most probably an underestimate.

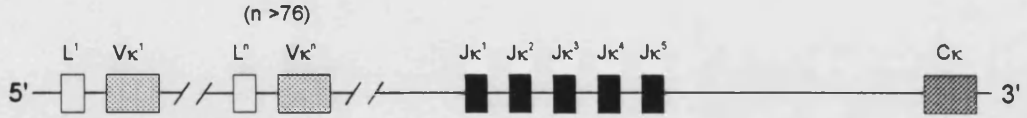
b - All seven constant lambda gene segments are identical.

Figure 1.4 - The genetic organization of the heavy, kappa and lambda loci in the human genome.
The letter L represents the leader sequence and symbols V, D, J and C depict the variable, diversity, joining and constant gene segments, respectively. Each CH gene is shown as a single box but in fact comprises of several exons. Asterisks indicate non-functional pseudogenes.

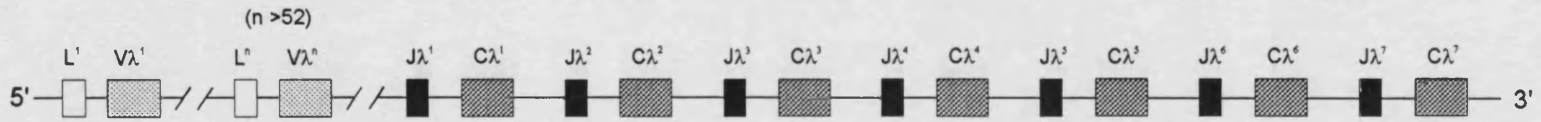
Heavy (H) chain locus (chromsome 14)



Kappa (κ) chain locus (chromsome 2)



Lambda (λ) chain locus (chromsome 22)



1.3.2 Rearrangement of immunoglobulin genes

Every clone of B-cells expresses a unique combination of V, D and J genes (Roitt *et al.*, 1986; Abbas *et al.*, 1991). The non-coding DNA between these DNA segments is removed and the coding regions are spliced together by a genetic mechanism known as DNA rearrangement. This site-specific recombination mechanism is achieved by an array of enzymes collectively termed recombinases. Specific recombination recognition sequences are found after all germline V, D, and J segments and are complementary to those preceding J and D segments (Tonegawa, 1983). Recognition of these complementary sites by recombinases is thought to allow the formation of DNA loop structures where the intervening DNA is excised and the cut ends are joined together. Joining of the gene segments is imprecise and contributes to antibody diversity i.e. an identical set of gene segments can result in different amino acid sequences (junctional diversity). The imprecise joining of the V-D-J segments in the heavy chain is particular important and creates additional diversity of the CDR3 loop. Extra genetic diversity of the heavy chain is also produced by insertion of nucleotides (N-sequences), not present in the immunoglobulin loci, at the rearranged D-J and V-D junctions. The heavy chain contributes most of the structural diversity to the antibody-combining site because of these extra genetic mechanisms.

The VDJ variable regions are initially expressed with a Cμ region (i.e. IgM) but during B-cell development the constant regions can be 'switched'. Upstream from each constant gene is a switch sequence (S) that allows the class (isotype) switching to occur via a 'DNA looping out' mechanism. This enables the B-cells to provide a variety of effector functions with the same antigen binding capacity.

1.3.3 The germinal centre reaction

T-cell dependent antibody responses initiate in the T-cell zones of the secondary lymphoid tissues. The majority of B-cells are activated in the T-cell zones migrate to local sites (the medulla cords and sinuses of the lymph nodes) where they differentiate into short lived plasma cells which produce fairly low affinity antibodies. Some of these B-cells bearing germline antigen

receptors (antibodies) with relatively high affinity for antigen are selected by T-cells and migrate to secondary follicles which contain follicular dendritic cell networks. Once the B-cells have expanded to fill the follicular dendritic cell network, compartmentalisation takes place and a germinal centre is formed. This all occurs within the first three weeks after antigen challenge and on average three B-cell clones colonise each follicle (reviewed in Maclennan, 1994; Kelsoe, 1996; Camacho *et al.*, 1998). There are a small number of antigens, which are capable of activating B-cells independently of T-cell help (T-independent antigens). These are large polymeric molecules and are unable to produce a secondary antibody response.

The germinal centre provides the location for both hypermutation and antigen selection of antibodies (a process known as affinity maturation) in a particular immune response. It involves the co-ordinated action of several different cell types: B-cells, T-cells, antigen-presenting cells, follicular dendritic cells and macrophages. The compartmentalisation of the germinal centre can be observed using a conventional light microscope and can be separated into two distinct sectors, the dark zone and the light zone (Figure 1.5).

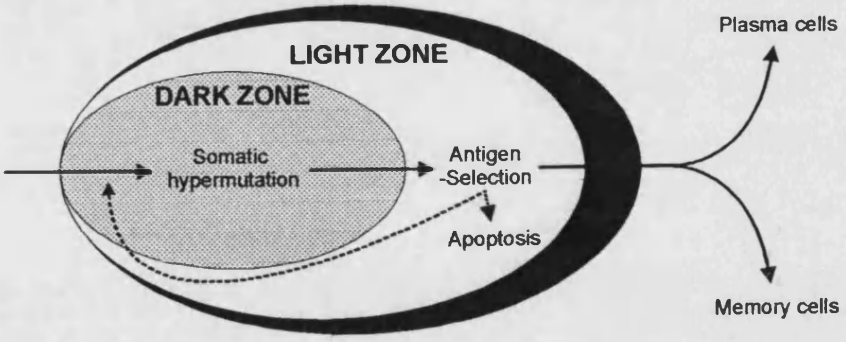


Figure 1.5 - Diagrammatic illustration of the germinal centre reaction.
The immunoglobulin DNA within the centroblasts is randomly mutated within the dark zone of the germinal centre. These centroblasts produce centrocytes, which express the mutated immunoglobulins on their cell surfaces. The centrocytes migrate to the light zone of the germinal centre and die by apoptosis unless they interact with the antigen. After a positive interaction with helper T-cells, the centrocytes differentiate into either memory or plasma B-cells. Additionally, the centrocytes can re-enter the dark zone (shown as a dashed line) and undergo further rounds of mutation and antigen-selection.

The dark zone is found adjacent to the T-cell zone in the secondary follicle and contains B-cells that are rapidly dividing and not expressing surface immunoglobulins (known as centroblasts). The somatic hypermutation (Section 1.3.4) machinery acts upon the variable gene segments of the centroblasts to produce randomly mutated DNA (Jacob *et al.*, 1991; Berek *et al.*, 1991; Pascual *et al.*, 1994). The centroblasts of the dark zone are continually giving rise to centrocytes which exit the cell cycle and display mutated immunoglobulins on their cell surfaces (Liu *et al.*, 1989). These centrocytes migrate to the light zone of the germinal centre where they are separated by a dense network of follicular dendritic cells, which display antigen on their cell surfaces for extensive periods of time (Section 1.3.5). Centrocytes are predisposed to apoptosis and survival depends on the ability of their surface receptors (mutated antibodies) to bind antigen held by the follicular dendritic cells (Liu *et* al., 1989; Liu and Arpin, 1997). One of the characteristics of apoptosis is the rapid fragmentation of DNA by a pre-existing endonuclease. Lindhout and colleagues (1985) showed that this DNA fragmentation in B-cells could be irreversibly blocked after interacting with follicular dendritic cells. The signals that prevent the B-cells from entering apoptosis are not fully characterised but the expression of oncogene, Bcl-2 has been shown to be important (Shokat and Goodnow, 1995). Apoptosis also occurs in the dark zone, which might reflect the death of centoblasts that have deleterious mutations such as stop codons in the immunoglobulin variable genes. Phagocytosis by tingible body macrophages is believed to provide the mechanism for eliminating dead or dying cells (Szakal *et al.*, 1989).

Once the specific interaction occurs between the antigen on the follicular dendritic cells and the antibody on the surface on the centrocytes, the antigen is internalised and processed by the centrocytes. These cells then re-express the antigen in a form that is recognised by the helper T-cells (Section 1.3.5). The specific interaction between the CD40 molecule on the surface of the B-cells (centrocytes) and the CD40-ligand on the surface of T-cells initiates isotype switching and differentiation of centrocytes into either plasma or memory cells. A positive CD40-mediated signal causes the B-cells to differentiate to memory cells whilst no signal produces plasma cells (Liu

and Arpin, 1997). Recent evidence indicates that follicular dendritic cells and T-cells within the germinal centre may not only support the proliferation of B-cells by also provide the signals necessary for activation of the hypermutation mechanism (Camacho *et al.*, 1998).

Three antigen-specific checkpoints in the germinal centre have been suggested to ensure no undesirable auto-reactive B-cells enter the memory B-cell pool (Lindhout *et al.*, 1997). The first checkpoint occurs outside of the germinal centre in the secondary lymphoid follicle where only antigen-specific B-cells are activated by T-cells to produce germinal centres. Secondly only centrocytes, which bind to the antigen on follicular dendritic cells with high affinity, will escape apoptosis. Finally, B-cells which are expressing irrelevant or auto-reactive antibodies will display the incorrect antigen on their cell surfaces, thus will fail to initiate the T-cell help required to differentiate into plasma or memory cells.

Theoretical studies suggest that the optimal mutation schedule for a germinal centre reaction is one in which the antigen-selected centrocytes of the light zone are recycled back into the dark zone for further rounds of mutation and selection (Kepler and Perelson, 1993; Oprea and Perelson, 1997). Recent experimental evidence shows the centroblasts of the dark zone are replenished from antigen-selected centrocytes of the light zone (Han *et al.*, 1995). Hence, it would seem that the compartmentalisation of the germinal centre with its cycles of mutation and selection provides a near optimal mutation schedule.

1.3.4 Somatic hypermutation

The precise molecular mechanism which mutates the antibody DNA of the centroblasts in the germinal centre reaction is unknown, but numerous basic features of the hypermutation process have been deduced from the sequences of thousands of antibodies recovered from different stages of the immune response (Kabat *et al.*, 1991).

The mutational process occurs in a non-random stepwise fashion and at an estimated rate of one mutation per thousand base pairs per generation (Berek and Milstein, 1988) which is a million times higher than the basal somatic mutation rate. The mutations are mainly nucleotide

substitutions and there is clear evidence of a strand bias (Neuberger and Milstein, 1995). Mutations only extend about 1 to 1.5 kilobases from the variable region promoter to the joining segments in the rearranged antibody DNA loci and are rarely found in the constant region. Low doses of antigen have been shown to produce a more focused set of mutations due to the effects of antigen selection on the germinal centre reaction (Gonzalez-Fernandez and Milstein, 1998).

Surprisingly, mutational 'hot-spots' exist which owe nothing to the skewing effect of antigenic selection (Betz *et al.*, 1993). In fact, the mutation machinery appears to intrinsically target sequence motifs in the CDRs of the variable gene which affect antigen binding compared to the framework regions which conserve the antibody structure (Tomlinson *et al.*, 1996: Green *et al.*, 1998).

Variable region genes have been replaced by numerous unrelated DNA sequences and hypermutation still occurs, hence showing the variable region genes themselves do not initiate the mutation process (Yélamos *et al.*, 1995). Numerous mechanisms for hypermutation have been suggested including an error-prone DNA polymerase or the lack of mismatch repair in DNA replication (Kocks and Rajewsky, 1989). Recent studies mostly support the idea that the transcriptional machinery plays an active role (Peters and Storb, 1996; Tumas-Brundage and Manser, 1997).

1.3.5 The 'alternative' antigen pathway

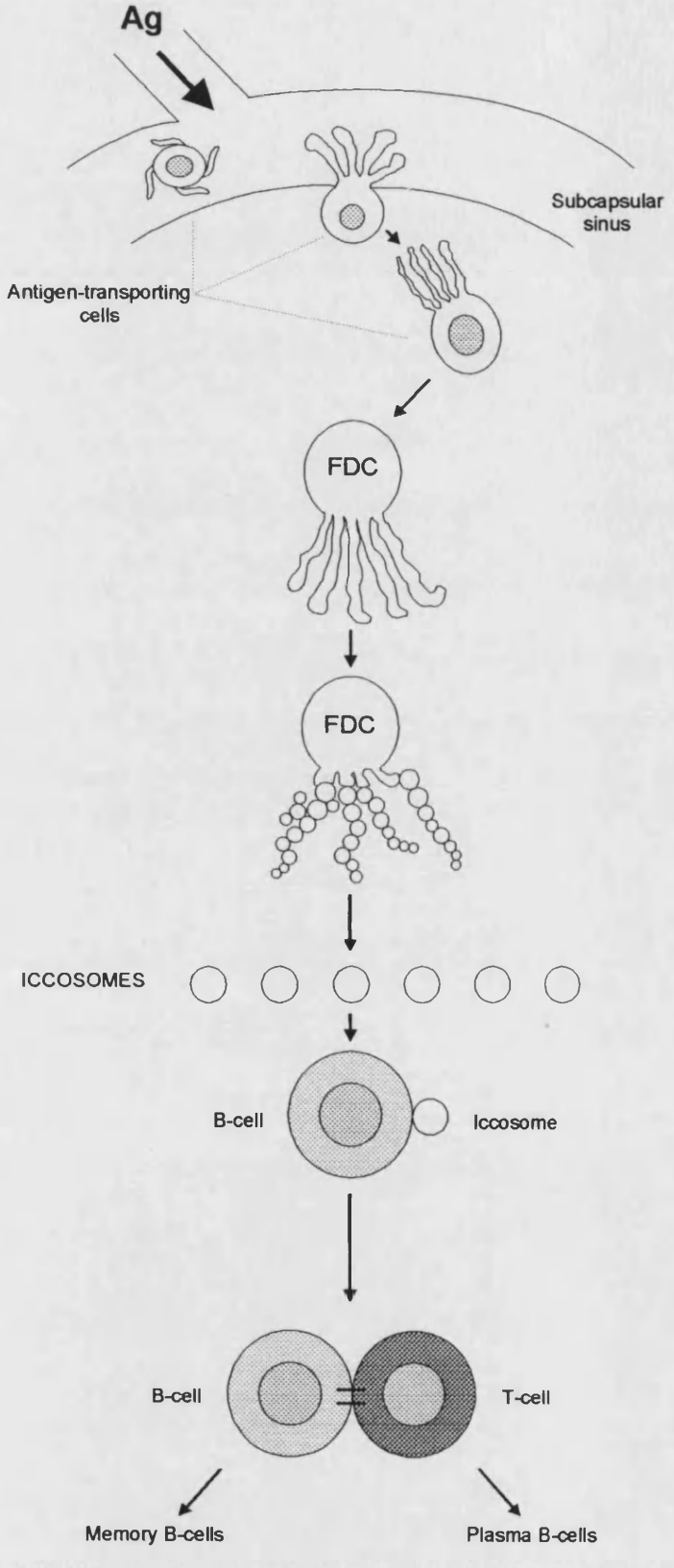
After booster immunisation, immune complexes with antigen rapidly form and enter one of two pathways. The majority of complexes are endocytosed, degraded and eliminated from the lymph nodes by macrophages [The 'classical' antigen pathway (Unanue, 1984)]

A second 'alternative' antigen pathway exists to preserve the immunogen (reviewed by Tew *et al.*,1989; Tew *et al.*,1997). Antigen transporting cells with dendritic morphology trap the immune complexes and migrates to the follicular dendritic cells present in the secondary lymphoid follicles. The antigen is then delivered to the follicular dendritic

cells by an unknown mechanism (Tew *et al.*, 1989). Scanning electron microscopy has revealed that the follicular dendritic cells retain the complexed antigen on the surfaces of their dendritic processes. Some of these dendritic processes mature and take on a 'beaded' morphology (Szakal *et al.*, 1988). The 'beaded' dendrites are named immune complex coated bodies (iccosomes), and have been shown to disperse and interact with B cells approximately three days after antigen challenge. The antigen presented by iccosomes is processed by the B-cells into endocytic vesicles, becomes associated with the Golgi lamellae and is re-expressed in a form recognisable to the antigen-specific helper T-cells (Lanzavecchia, 1985; Kosco *et* al., 1988; Szakal *et al.*, 1988). If helper T-cells recognise the antigen presented by the surface of B-cells they produce co-stimulatory signals such as lymphokines necessary for further B-cell differentiation. This alternative antigen pathway is described in figure 1.6 and follows the so-called antigen transport-follicular dendritic cell-iccosome-B cell axis.

Some of the antigen on the surface of the follicular dendritic cells does not take part in iccosome formation and is retained within the dendritic processes like 'balls of yarn' for long periods of time. This retained antigen is thought to play an important role in the regulation and long-term maintenance of humoral immunity. When antibody levels are high, newly produced antibody covers the antigen exposed on the surface of the follicular dendritic cells and consequently terminates the immunogenic stimulus. Cross-linking of the antigen by antibodies on neighbouring dendrites result in the formation of the 'ball of yarn' morphology. Antigen is then hidden in the convolutions of the dendrites until the antibody levels decline and the antigen becomes exposed to stimulate memory B-cells (Szakal *et al.*, 1989). This cycle of events may explain how the levels of a particular antibody are maintained for months or sometimes years.

Figure 1.6 - The alternative antigen pathway.



1) *Antigen transport* (1min-24hrs).
Immune complexes with antigen form rapidly in the subcapsular sinus and are delivered to FDCs at the peripheral of the germinal centre.

2) *Iccosome formation and dispersion* (Days 1-3).
Iccosomes form on FDCs, disperse and interact with B-cells.

3) *Iccosome endocytosis* (Day 5).
The antigen presented by iccosomes is endocytosed and re-expressed on the surface of B-cells in a form recognizable by helper T-cells.

4) Ag processing and presentation to helper T-cells (Days 3-7).
Co-stimulatory signals from helper T-cells stimulate further differentiation of B-cells into plasma or memory cells.

Note: Steps 2 - 4 occur within the germinal centre. After step 4, antigen is retained on the FDCs for long periods of time. Once antibody levels decline, the antigen is exposed and memory B-cells are stimulated. Abbreviations: Ag, Antigen; FDC, follicular dendritic cell; Ab, antibody; iccosomes, immune complex-coated bodies.

1.4 ANTIBODY ENGINEERING

1.4.1 Production of polyclonal and monoclonal antibodies

Immunisation of a mouse with protein antigen A initially produces a variety of antibodies that are relatively low in affinity. Over the course of the response (repeat immunisation), the affinity of the antibodies to the protein antigen A increases due to affinity maturation (Section 1.3.3). The antiserum from the blood of this mouse will contain a variety of antibodies to different epitopes carried on the protein antigen A, all produced by different plasma cells (Polyclonal response). Additionally, this antiserum will comprise of a set of different antibodies with no affinity for protein antigen A. Once the serum from this mouse is used, it is impossible to produce the same set of antibodies by immunisation of a different mouse. These disadvantages of immunisation can be overcome by using monoclonal antibodies (an homogeneous population of antibodies with the same specificity).

Köhler and Milstein in 1975 first described the production of monoclonal antibodies using hybridoma technology. Hybridomas can be produced by fusing myeloma cells (cancerous B-cells) to antibody-producing spleen cells (or lymph node cells) in the presence of polyethylene glycol. Selective media can then be used to select for hybrid cells which have retained the immortal characteristic of the myeloma cells and the antibody-producing capacity of the spleen cells. Cell lines generating antibodies with single specificity for the desired antigen can then be isolated directly.

This powerful hybridoma technology has been applied successfully to a wide variety of biological problems, including many diagnostic applications. Unfortunately, the production of human monoclonal antibodies using this technology has not been so spectacular. This is mainly due to technical problems such as the instability of human heterohybrids, the loss of relevant structural/regulatory genes and imperfections in the secretory machinery of cells (James and Bell, 1987). Furthermore, the ethical problems associated with obtaining a good source of human B-cells from either the spleen or lymph nodes have hindered progression. The difficulty in

making human monoclonal antibodies, especially against human self-antigens, has prompted the development of human antibodies using genetic engineering techniques.

1.4.2 Humanisation of antibodies

Rodent monoclonal antibodies offer a plentiful and pure source of effective antibody reagents but if used *in vivo* they are recognised as foreign (i.e. produce a human anti-mouse immune response) and fail to recruit effector functions. Genetic and protein engineering techniques have provided the tools to overcome these problems by a process known as 'humanisation'.

The first generation of humanised antibodies consisted of rodent variable domains being transplanted onto the constant domains of a human antibody (Morrison *et al.*, 1984). These engineered fragments were shown to reduce immunogenicity but a residual immune response against the rodent variable domains decreased its effectiveness for clinical uses.

Due to the conservation of antibody structures across species, the humanisation process was further developed by transplanting the rodent CDRs onto a human β-sheet framework (Jones *et al.*, 1986; Verhoeyen *et al.*, 1988). If the framework-CDR antibody interface is accurately reconstructed during this grafting process, the antigen-combining site is conserved and affinities close to the parental rodent antibodies are achievable (Holmes and Foote, 1997). Often additional substitutions of individual amino-acid residues within the framework are necessary to conserve the rodent antigen-combining site (Foote and Winter, 1992). Computational methods have been developed to determine which human framework is suitably similar to the rodent framework to conserve CDR loop structures and minimise the number of foreign residues (Queen *et al.*, 1989).

A novel technique known as resurfacing has also been developed to reduce the immunogenicity of the murine antibodies by humanising the surface accessible residues located at the conserved positions in the rodent's framework (Roguska *et al.*, 1996).

1.4.3 Engineered antibody fragments

The domain structure of the IgG molecule makes it particularly accessible to protein engineering and can be separated into functional domains carrying either the antigen-binding (Fab) or effector (Fc) capacities by limited proteolysis. Digestion with the protease pepsin removes the second and third constant regions (CH2 and CH3) leaving the two Fab fragments joined together by at least one disulphide bond [Figure 1.7 ($F(ab')_2$)]. Using another protease known as papain can also produce individual Fab fragments (Figure 1.7).

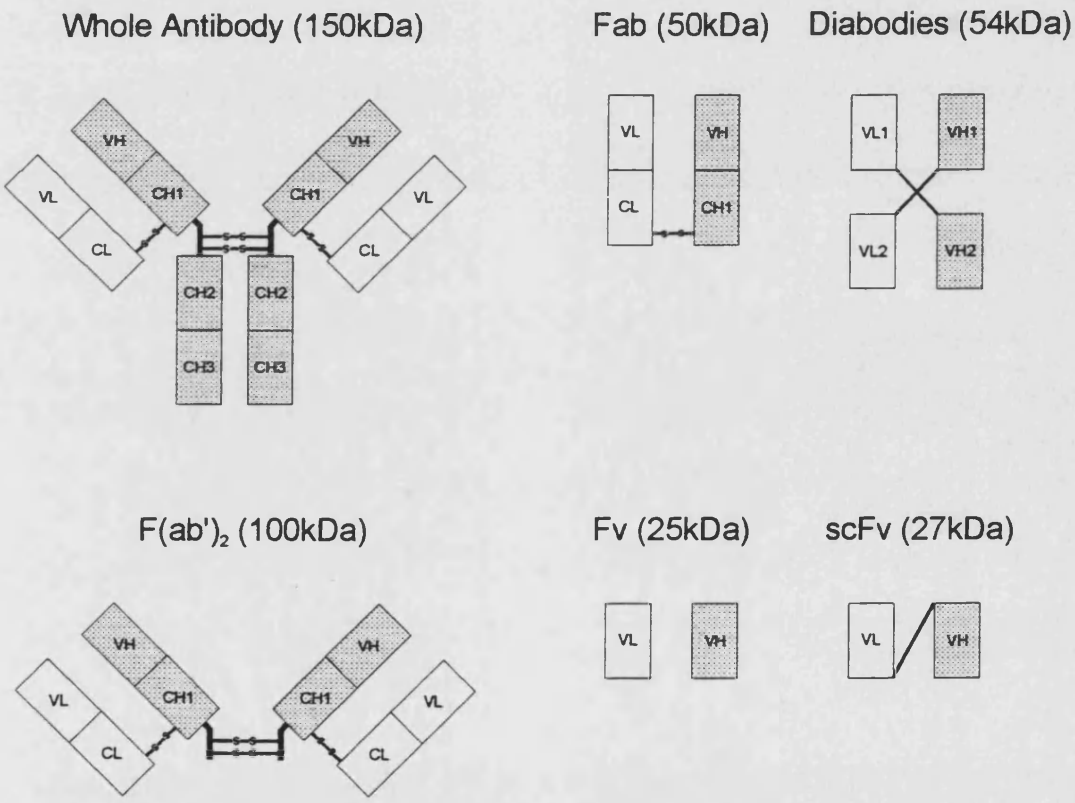


Figure 1.7 - Engineered antibody fragments.

The various formats of antibody engineered fragments compared with the native immunoglobulin. The whole antibody molecule can be split into one $F(ab')_2$ fragment which consists of the heavy (VH + CH1) and the light chain variable and constant regions (VL + CL) held together by at least one disulphide bond. The $F(ab')_2$ molecule can be separated into two functional antigen-binding fragments (Fab). When the CL and CH1 regions are removed, the remaining variable regions make up the Fv fragment. Single-chain Fvs (scFvs) have the heavy and light chains linked by a short peptide linker. Diabodies are formed by shortening the linker of the scFv fragment to force dimerisation.

The smallest antibody fragment which contains the entire antigen-binding site is the fragment variable (Fv). This Fv consists of the variable heavy and light antibody chains held together by non-covalent bonds which tend to dissociate at low concentrations (Glockshuber *et al.*, 1990). The advent of the single chain Fv [scFv (Figure 1.7)] which joins these two variable regions by a peptide linker overcame this dissociation problem (Bird *et al.*, 1988; Huston *et al.*, 1988).

Comparisons of the scFv's affinities with the parental antibodies have shown the scFv affinities to be slightly lower but essentially the same (Bird and Walker, 1991). Normally, the orientation of the variable regions in the scFv (i.e. the linker running from the C-terminus of the VH to the N-terminus of the VL or vice versa) does not affect affinity, unless the C-terminus or N-terminus of either of the variable domains contribute to the antigen-binding site (Desplancq *et al.*, 1994).

The peptide sequence spanning the distance between the variable domains in the scFv must be sufficiently long and flexible enough to allow the formation of the antigen-combining site (Johnson and Bird, 1991). Molecular modelling based on structural data has been used to successfully select for natural linkers with the correct molecular dimensions and length to span the distance between the variable domains (Bird *et al.*, 1988). In addition to length, amino acid composition plays a crucial role in linker design and minimal secondary structure is desirable to inhibit any interference with the amino acids of the variable domains (Huston *et al.*, 1988; Argos, 1990). Commonly, linkers between 14 and 25 residues in length and various amino acid sequences are used (Huston *et al.*, 1991). The $(Gly_4Ser)_3$ linker used in this project is designed with glycine residues which provides the necessary flexibility and hydrophilic serine residues which promotes stability by the formation of hydrogen bonds with the surrounding solution (Huston *et al.*,1988). Often, the C- and N-termini of the variable domains are flexible enough to serve as part of the linker if the peptide is too short.

Antibody fragments have particular advantages in tumour therapy and imaging due to their small size (rapid pharmokinetics and tumour

penetration *in vivo*) and the absence of the side reactions associated with the Fc effector function (reviewed in Chester and Hawkins, 1995). Therapeutic applications using antibody fragments to date include *in vivo* diagnosis of cancer using radiolabels (Schlom *et al.*, 1992; King *et al.*, 1994; Neri *et al.*, 1997), antibody fusions with cytotoxic agents [immunotoxins (Ghetie and Vitetta, 1994)] and antibody-directed enzyme prodrug therapy (Bagshawe, 1989). Additionally, antibody fragments are ideal models for protein engineering studies and have been used extensively to investigate various antibody-antigen interactions.

ScFvs with short linkers have a tendency to form dimers and multimers (Desplancq *et al.*, 1994; Alfthan *et al.*, 1995). These dimers arise from a head-to-tail intermolecular association of the variable heavy and light domains due to the short linkers inhibiting the folding needed for the intramolecular event (Figure 1.7). This property can be used to produce 'diabodies' where short linkers are designed to force pairing of complementary domains on opposite chains to create two antigen combining sites (Holliger *et al.*, 1993). Fab and scFv antibody fragments lack the region of the antibody that recruits natural effector functions (Fc fragment), therefore diabodies have been developed to overcome this disadvantage. Directing one antigen-combining site towards the antigen and the other towards an effector receptor such as the first component of the complement system [Clq (Kontermann *et al.*, 1997)] or a natural immunoglobulin molecule (Holliger *et al.*, 1997) have been shown to be fruitful. The clinical data obtained from treatment of tumours with various diabodies is particularly promising (van de Winkel *et al.*, 1997).

In 1993, a new class of antibodies was discovered in the camelid species that lacked light chains (Hamers-Casterman *et al.*, 1993). This camelid variable heavy domain replaced the Fv as the 'minimal antigen-binding fragment' (Sheriff and Constantine, 1996) and are expected to become extremely useful in a number of biotechnological and medical applications (Muyldermans and Lauwereys, 1999).

1.4.4 Expression of recombinant antibody fragments

Until the advent of the polymerase chain reaction (Saiki *et al.*, 1985), the cloning of antibody genes was a laborious and time-consuming process involving the creation and screening of genomic or cDNA libraries. The polymerase chain reaction has made it possible to reproduce antibody gene repertoires present *in vivo* (Orlandi *et al*, 1989). 'Universal' primers can be designed to the conserved regions of the 5' and 3' ends of the variable heavy and light DNA sequences present in the Kabat database. To date, this strategy has been used to prepare antibody gene repertories from hybridomas (Orlandi *et al.*, 1989), mouse spleens (Huse *et al.*, 1989; Ward *et al.*, 1989) and peripheral blood lymphocytes (Marks *et al.*, 1991).

Most laboratories generally express their chosen antibody gene fragments in *Escherichia coli* because of the relatively simple, fast and well-established methods available (reviewed in Plückthun, 1991; Ward, 1992). Initial attempts to express antibodies intracellularly in *E. coli* produced very low yields of functional antibody (Cabilly *et al.*, 1984; Boss *et al.*, 1984). The expressed protein was contained in 'inclusion bodies' and the generation of functional antibody required *in vitro* denaturation and refolding. A significant breakthrough was achieved when signal sequences were used to target antibody expression to the periplasmic space (Better *et al.*, 1988; Skerra and Plückthun, 1988). The oxidising environment of the periplasmic space promotes the formation of disulphide bonds, hence functional antibody can be isolated directly without any further manipulations.

E. coli production of antibody fragments is not always straightforward and unpredictable expression biases exist (Knappik and Plückthun,1995). Eukaryotic systems such as plants (Owen *et al.*, 1992) and mammalian cells (Jost *et al.*, 1994), which have more sophisticated folding pathways, have also been developed for antibody fragment expression. One of the most successful technologies used in the development of specific antibody fragments is phage display.

Phage display

Understanding the way antibodies are made *in vivo* has inspired the creation of an artificial immune system using phage display technology (reviewed in Winter *et al.*, 1994; Winter, 1998). This technology has been successfully used to develop numerous antibody fragments with different binding specificities for both foreign and self-antigens, often with such efficiency that hydridoma technology and immunisation can be surpassed.

To mimic the immune system, firstly a repertoire of antibody genes must be created. 'Natural' repertories of rearranged human variable genes from peripheral blood lymphocytes of unimmunized donors (Marks *et al.*,1991) and 'synthetic' repertories of unrearranged variable gene segments with randomised CDR3 (Hoogenboom and Winter, 1992; Griffiths *et al.*, 1994) have been produced.

The 'genetic display package' of the B-cell is copied by displaying these antibody libraries on the surface of a filamentous bacteriophage using the molecular information of its coat proteins. Valencies of 3-5 or >24 copies of antibodies per phage are achieved by the fusion of the antibody DNA to the minor [pIII (McCafferty *et al.*, 1990) or the major [pVIII (Kang *et al.*, 1991)] coat proteins, respectively. Both Fab (Hoogenboom *et al.*, 1991) and scFv (McCafferty *et al.*, 1990) antibody fragments have been successfully displayed on the surface of phage.

The antigen-driven selection of the immune response is imitated by panning of the phage display libraries with different types of antigen coated surfaces. Several rounds of selection can isolate a single phage with high affinity for the pre-defined antigen. Additionally, it is possible to take advantage of the differing dissociation rates (off-rates) of the phage from the antigenic surface to isolate antibodies with closely related affinities but differing off-rates (Hawkins *et al.*, 1992). Multivalent interactions also improve discrimination between phages i.e. small differences in affinity can give rise to larger differences in avidity.

It is also possible to copy the function of antibody-producing plasma cells by placing an amber stop codon in-between the antibody and coat

protein DNA sequences. When phage is grown in a suppressor strain of *E. coli* (supE), the amber stop codon reads as a glutamine and the antibody fusion is displayed on the surface of phage (analogous to memory cells). When phage is grown in a non-supressor strain, the amber codon is read as a stop codon and soluble protein is secreted into the periplasm (analogous to plasma cells).

Previously, the size of a phage antibody library was limited to $<10^9$ clones by the efficiency of the transfection of DNA into bacteria. Waterhouse and coworkers (1993) overcame this problem by a novel technique known as 'combinational infection and *in vivo* recombination' and a library of 6.5 x 10^{10} colonies was obtained (Griffiths *et al.*, 1994). In this method, bacteria harbouring a 'donor' heavy chain repertoire (on a plasmid) was infected by an 'acceptor' light chain repertoire (on phage) and then Cre-catalysed site-specific recombination was used to combine the heavy and light chains. Theoretically the larger the library, the greater the chance of finding a higher affinity antibody for any given epitope (Perelson and Oster, 1979). Isolation of high affinity antibodies from very large antibody libraries (Griffiths *et al.*, 1994; Vaughan *et al.*, 1996) supports this hypothesis and antibodies to any given antigen with affinities comparable with those from a secondary immune response have been isolated directly.

Additionally, antibody fragments have been expressed on the external surface of *E. coli* by using outer membrane proteins such as the major lipoprotein-outer membrane protein A hybrid [Lpp-OmpA (Francisco *et al.*, 1993)] and peptidoglycan associated lipoprotein [PAL (Fuchs *et al.*, 1991)] display systems. These have been successfully used with fluorescence-activated cell sorting (FACS) to directly isolate single *E. coli* cells displaying specific antibodies.

An antigen-specific selection process has also been developed which utilizes both bacterial and phage display systems (Malmborg *et al.*, 1997). Phage infection is reliant on the minor coat protein (pIII) on the bacteriophage interacting with the F pilus on the bacteria. Bacteriophages were designed to display peptide-specific antibody fragments using the pIII

coat protein and bacteria were engineered to express a particular antigenic peptide using a protein subunit of F pilus (*traA*). Specific interaction between the antibody fragment and peptide resulted in bacterial infection.

Expression using the yeast Saccharomyces cerevisiae

Expression of antibodies in *S. cerevisiae* was first reported by Wood and co-workers in 1985. Although functional Fab antibody fragments were formed, the efficiency of assembly was very low and the antibodies accumulated in the vacuoles. High intracellular levels of functional catalytic antibody fragments were achieved using an ultrahigh copy number plasmid and a galactose-inducible promoter (Bowdish *et al.*, 1991). The expressed catalytic antibody possessed chorismate mutase activity and was shown to be active in an engineered *S. cerevisiae* strain which lacked this functional enzyme (Bowdish *et al.*, 1991; Tang *et al.*, 1991). Additionally, antibody fragments have been expressed intracellularly to eliminate the function of a single enzyme (Carlson, 1988).

The first secreted Fab antibody fragment in *S. cerevisiae* was achieved by simultaneous expression of the light and heavy chains fused to the yeast invertase signal sequence (Horwitz *et al.*, 1988). Although fully functional antibody fragments were formed, the efficiency of assembly was low. Recently, secretion of single chain Fv antibody fragments in *S. cerevisiae* has reached levels up to 20mg/ml by using a combination of expression level tuning and overexpression of folding assistants (Shusta *et al.*, 1998). This project is concerned with the expression of antibody fragments (Fab and scFv) on the surface of *S. cerevisiae*.

1.4.5 Antibody improvement *in vitro*

For clinical applications, antibodies with high affinities (Ka ~10^9 M^{-1}) are required to bind tightly and specifically to a particular therapeutic target (Schlom *et al.*, 1992). Antibodies with high affinities can be isolated directly from very large antibody repertoires (Section 1.4.3) but antibodies selected from medium size naive, synthetic or immune phage libraries normally reflect

the affinities typical of primary immune responses (Ka ~10^5 -10^7 M^{-1}) and require improvement (Hoogenboom, 1997).

In vivo, antibody affinities are improved in the germinal centre by a step-wise process of mutation and selection (Section 1.3.3). This process has been imitated *in vitro* by adding mutations to the antibody variable regions using either an *E. coli* mutator strain (Low *et al.*, 1996) or error-prone PCR (Hawkins *et al.*, 1993) and selecting for improvements using phage display. Antibodies with high affinities for a number of therapeutic targets have been isolated using these methods.

Although mutations in framework residues may contribute to improved affinity, many successful techniques have focused on changing the CDRs of the antibody, which have a major influence on antigen binding. CDRs have been targeted using numerous mutatgenic techniques including using specific oligonucleotides engineered to reduce the number of amino acids which do not retain parental structural features [parsimonious mutagenesis (Schier *et al.*, 1996)] and an oligonucleotide-directed method that replaces entire codons rather than individual nucleotides [codon-based mutagenesis (Yelton *et al.*, 1995)]. A combination of CDR mutagenesis and selection of antibodies with improved affinity (CDR walking) has been successfully used to generate antibodies with picomolar affinities, especially when individually selected variants are amalgamated (Barbas *et al.*, 1994; Yang *et al.*, 1995).

A successful alternative to random point mutagenesis is a process known as chain shuffling. Three hundred-fold antibody improvements have been achieved by pairing a combination of heavy and light chains with the appropriate single partners (Clackson *et al.*, 1991; Marks *et al.*, 1992).

Antibody fragments that have been improved *in vitro* are now in clinical trials in areas such as tumour imaging and targeting (Neri *et al.*, 1997).

1.5 THE CHOSEN EXPRESSION HOST - *Saccharomyces cerevisiae*

Saccharomyces cerevisiae is a unicellular eukaryote that can exist in any of three different cell types- **a**, α and **a**/α. Its life cycle can be divided into two parts: cellular proliferation and change in ploidy. All three cell-types can mitotically proliferate but only **a**- and α-haploids cells can mate to form the diploid, **a**/α. Upon nutritional starvation, diploids are capable of undergoing meiosis which gives rise to four haploid cells (Herskowitz, 1988).

The secretory system in *S. cerevisae* is very similar to higher eukaryotes, therefore it is very useful for studying many of the basic questions of eukaryotic biology including gene expression, DNA replication, secretion and the cell cycle. Properties such as a stable haploid and diploid state, availability of good genetic markers, quick growth, simple inexpensive media and a small haploid genome [1.4 x 10^7 b.p. (Sherman, 1991)] have made *S. cerevisiae* particularly suitable for genetic studies. Replicating plasmids are available with origins of replication and selectable markers for both *E. coli* and *S. cerevisiae.* Therefore, *E. coli* can be used as a workhorse for all the genetic manipulations and *S. cerevisiae* as the expression host. Additionally, homologous recombination is very efficient in *S. cerevisiae* and marker genes can be directly integrated into specific locations in the genome, thus permitting simple ways to study gene function.

With the completion of the total DNA sequence of the genome of *S. cerevisiae* in the spring of 1996 (Goffeau *et al.*, 1997), new opportunities have developed to study gene function. Computer analysis of the genome sequence of the sixteen yeast chromosomes has identified more than six thousand potential open reading frames. Thirty-five percent of these potential genes have no known function and are the subject to substantial international efforts such as the European Functional Analysis Network (EUROFAN) which is a network of European laboratories studying the function of these genes by gene disruption. The total DNA sequence of *S. cerevisiae,* including the mitochondrial genome (Foury *et al.*, 1998), is

available in the *S. cerevisiae* genome database (SGD) on the World Wide Web.

As well as *S. cerevisiae* having very well established genetics, it is also an ideal host for the expression of eukaryotic proteins (Romanos *et al.*, 1992; Hammond and Helenius, 1995). Its secretory system is similar to mammalian cells; hence it is able to undertake all post-translational modifications such as glycolysation and disulphide bond formation. It expresses more protein per litre than mammalian systems (but less than *E. coli*) and is a food organism, therefore has 'generally regarded as safe' (GRAS) status.

All the above properties make *S. cerevisiae* highly acceptable for the production of pharmaceutical proteins and numerous commercial successes including insulin (Thim *et al.*, 1986) and hepatitis-B surface antigen (Valenzuela *et al.*, 1982) have been reported. *S. cerevisiae* is the chosen expression system for this project and factors which can affect expression levels are described in Part One. Additionally, the mating characteristics of *S. cerevisiae* were utilized in the second part of this project (Chapters 5 and 6) and therefore are reviewed in the next Section.

1.6 THE MATING PROCESS IN *Saccharomyces cerevisiae*

1.6.1 Pheromone induction

The first stage in the mating process is communication between the two haploid cell types, **a** and α. This is achieved via the cell-type specific production of secreted mating factors (pheromones). The *MAT*α cell-type secretes a 13-residue pheromone, α-factor and expresses a cell surface receptor that binds **a**-factor. **a**-factor (a 12-residue peptide) is secreted by the *MAT***a** cell type and it also expresses a cell surface receptor for recognition of α-factor (Herskowitz, 1988). Both **a**- and α-factor precursor are encoded by two functional genes *MF**a**1/MF**a**2* and *MFα1/MFα2,* respectively (Michaelis and Herskowitz, 1988, Singh *et al.*, 1983).

Mutual stimulation of the mating pair by pheromones activates a common response pathway which causes arrest in cell division (G1 phase of the cell cycle), increased agglutinability and morphological changes which cause the cells to elongate towards each other to form pear-shaped cells [sometimes called shmoos (Bardwell *et al.*, 1994)]. These processes synchronise the mating partners for both cell and nuclear fusion. Haploid cells that do not mate successfully are able to recover from pheromone-imposed cell cycle arrest and resume vegetative growth.

1.6.2 Agglutination

Cellular aggregation (or agglutination) occurs through complementary binding of the mating type specific glycoproteins, **a**- and α-agglutinin. This interaction mediates adhesion between cells of the opposite mating types during the mating process. Agglutinins are constitutively expressed at a low level on the cell surface of haploid cells (< 1 x 10^3 agglutinin molecules per cell) but expression is strongly enhanced by mating pheromones from the opposite phenotype [3-5 x 10^4 agglutinin molecules per cell (Terrance and Lipke, 1987; Watzele *et al.*, 1988; Wojciechowicz and Lipke, 1989). Initiation of agglutination takes approximately twenty minutes and the agglutinins are often found concentrated at the tips of shmoos (Watzele *et al.*, 1988).

Agglutinins are essential for mating under conditions that do not promote cell to cell contact (liquid medium) but are not essential under conditions that promote cell to cell adhesion (solid medium). Hence, all agglutinin mutants have an unusual mating phenotype where the mating frequencies are only slightly affected on solid medium but when assayed on liquid medium the frequencies decrease by several orders of magnitude compared to the wild-type isogenic strains (Lipke *et al.*, 1989; Roy *et al.*, 1991). This phenotypic characteristic has been utilised to identify the agglutinin structural genes via two different approaches:

1) A genetic approach - agglutination mutants which only affect one mating type are isolated and the agglutinin genes are cloned by complementation.

2) A biochemical approach - the relevant agglutinin is purified, the amino acid sequence is obtained and degenerate oligonucleotides corresponding to these sequences are used to clone the agglutinin genes.

α-Agglutinin

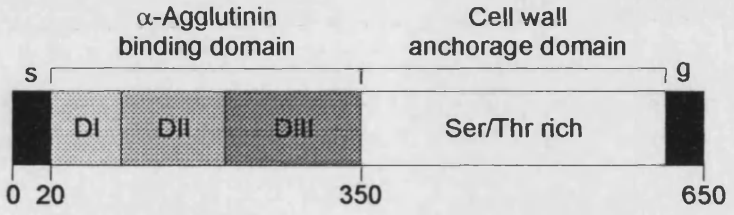
The α-agglutinin gene sequence (*SAG1*; Doi *et al.*, 1989) was determined both biochemically (Hauser and Tanner, 1989) and genetically (Lipke *et al.*, 1989). The DNA sequence obtained (1950 b.p.) encodes for a 650 amino acid open-reading frame containing a nineteen amino acid leader sequence and twelve potential N-glycosylation sites. The mature α-agglutinin protein (Sag1p) is highly N-glycosylated (~50% carbohydrate) and has a molecular weight of over 200 kDa but the total removal of these carbohydrates does not effect the biological activity of the protein (Hauser and Tanner, 1989).

Analysis of truncated versions of the α-agglutinin structural gene revealed it consists of two functional domains: an N-terminal half which binds the **a**-agglutinin and a C-terminal half which is important for cell surface attachment (Wojciechowicz *et al.*, 1993). The C-terminal half (amino acids 350-650) comprises of approximately 50% serine and threonine residues and a signal sequence for the attachment of the glycosyl phosphatidylinositol (GPI) anchor (Figure 1.8). The N-terminal part of the α-agglutinin protein (residues 20-350) is composed of three immunoglobulin variable-like domains (Wojciechowicz *et al.*, 1993; Chen *et al.*, 1995) and shares similarities with many of the features of eukaryotic cell adhesion proteins. A histidine residue at position 292 is critical in the **a**-/α-agglutinin interaction (Cappellaro *et al.*, 1991). Further mutagenesis studies identified a putative binding site comprising of three discontinuous patches in domain III of the α-agglutinin protein (de Nobel *et al.*, 1996).

Figure 1.8 - Diagrammatic representation of the **a**- and α-agglutinin structural features.

α-agglutinin

The α-agglutinin consists of an N-terminal binding domain (residues 20-350) and a C-terminal anchorage domain (residues 350-650). It is proposed that the N-terminal α-agglutinin binding domain consists of three immunoglobulin variable like domains called DI, DII, and DIII and a putative binding site has been identified in domain III.



a-agglutinin

The **a**-agglutinin consists of two subunits; an anchorage subunit (AGA1p) and a small binding subunit (AGA2p) which are held together by two disulphide bonds.

All three agglutinins contain a signal sequence (s) that targets the protein to the secretory pathway. Both AGA1p and SAG1p contain a C-terminal signal sequence (g) for the attachment of the glycosyl phosphatidylinositol (GPI) anchor.

a-Agglutinin

Genetic analysis of **a**-agglutinin revealed it consists of two glycoprotein subunits which are encoded by genes *AGA1* and *AGA2*. The *AGA1* gene (2175 b.p.) was cloned by complementation (Roy *et al.*, 1991) and analysis indicated the protein contains a large amount of serine/threonine residues and a C-terminal signal sequence for attachment of the GPI anchor (Figure 1.8). Evaluation of *aga1* mutants confirms that it is involved in cell surface attachment of the **a**-agglutinin (Roy *et al.*, 1991).

Treatment of **a**-cells with reducing agents eliminates cellular agglutination due to the release of the **a**-agglutinin binding subunit. The open-reading frame of the **a**-agglutinin binding subunit (*AGA2*) was determined biochemically (Cappellaro *et al.*, 1991) and it encodes a 69 amino acid protein which contains an N-terminal signal sequence and is highly O-glycosylated (Figure 1.8). The **a**-agglutinin binding subunit (Aga2p) binds to the **a**-agglutinin cell anchorage subunit (Aga1p) via two disulphide bridges. The binding activity of **a**-agglutinin is located in the last ten amino acids of the C-terminus of Aga2p (Cappellaro *et al.*, 1994).

The **a**- and α-agglutinins interact on the cell surface in a monovalent form (Cappellaro *et al.*, 1991). The interaction is heterogeneous and both weak ($Ka<10^8 M^{-1}$) and strong ($Ka = 1 \times 10^9 M^{-1}$) association constants have been reported (Lipke *et al.*, 1987). The strength of the agglutination reaction is more closely related to the number of agglutinin molecules on the surface of *S. cerevisiae* than the binding constant (Lipke *et al.*, 1987).

The secretory pathway

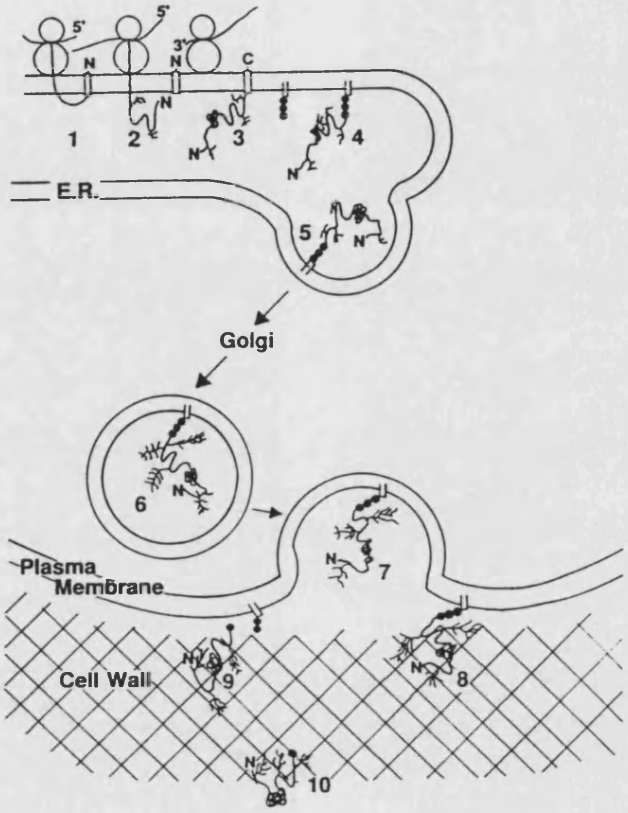
All three agglutinin structural genes contain N-terminal hydrophobic regions. These are likely to act as signal sequences which direct the agglutinins through the secretory pathway of *S. cerevisiae* [Figure 1.9 (Tohoyama and Yanagishima, 1985; Lu *et al.*, 1994)].

After the agglutinins have been directed through the secretory pathway the anchorage domains (SAG1p and AGA1p) become covalently attached to the cell wall. The cell wall of *S. cerevisiae* consists of two major components: β-glucans (β-1,3/β-1,6-glucan) and mannoproteins. It is a layered structure and its composition varies considerably throughout the cell cycle. The outer surface of the cell wall consists predominantly of mannoproteins, has a brush-like structure and is often named the fibrillar layer. The inner layer mainly consists of β-glucans and these are responsible for the mechanical strength of the cell wall (Klis, 1994).

Both anchorage domains (SAG1p and AGA1p) contain C-terminal hydrophobic signal sequences for the attachment of the GPI anchor that is

essential for covalent linkage to the β-1,6-glucan component of the cell wall (Roy *et al.*, 1991; Lipke *et al.*, 1989). The mechanism of cell wall attachment is complex and relies on the prior addition of the GPI anchor (Lu *et al.*, 1995). A recent study has shown that only a short amino acid sequence (ω-minus) in the GPI anchor is important for localization to the cell wall (Hamada *et al.*, 1998). It has been proposed that the periplasm form of the agglutinin is cross-linked to preformed β-1,6-glucan which becomes integrated into the growing cell wall (Lu *et al.*, 1995). It is speculated the agglutinins interact within the cell wall's outer fibrillar layer (Cappellaro *et al.*, 1994).

Figure 1.9 - The secretory pathway of *S. cerevisiae.*



The signal sequence mediates translocation of the agglutinin molecules into the endoplasmic reticulum (ER). The signal peptide is removed by a signal peptidase in the ER and O- and core N-linked glycosylation occur [attachment of the GPI anchor also happens in the ER (Step 1-5)]. The agglutinins are then transported in vesicles to the Golgi where modifications of these glycosyl structures take place (not shown). From the Golgi, the agglutinins are packaged into secretory vesicles (Step 6) and delivered to the cell surface (Step 7). The agglutinins are then transferred from the GPI anchor to the cell wall glycan (Step 8-9). The mature agglutinins remain anchored to the cell wall with the binding domain exposed at the cell surface (Step 10). Reproduced from Lipke and Kurjan, 1992.

Regulation of expression

Agglutinin expression in *S. cerevisiae* is specific to its haploid mating-type and is subject to a variety of types of regulation including the mating-type locus, induction by pheromones and growth conditions.

Genetically, **a**- and α-haploid cells only differ at the single mating-type locus, *MAT*. The expression of the agglutinins is regulated at the *MAT* locus by a mechanism similar to other **a**- and α-specific products. The *MATα* locus encodes two DNA binding proteins: MATα1p, which is a positive regulator of α-specific genes and MATα2p, which is a negative regulator of **a**-specific genes. α-Agglutinin expression is positively regulated by MATα1p and the MATα1p consensus sequence is found upstream of the *SAG1* coding region. In **a**-cells, the **a**-specific genes are constitutively expressed and MATα2p negatively regulates Aga2p expression in α-cells (Lipke and Kurjan, 1992). Surprisingly, the *AGA1* transcript is expressed in both **a**- and α-cells at low levels and the mating-type control of the **a**-agglutinin binding subunit (Aga2p) ensures that the active protein is only expressed in **a**-cells (Roy *et al.*, 1991).

Both the *SAG1* and *AGA2* transcripts are also induced by the opposite mating-type pheromone and contain pheromone induction consensus sequences upstream from their start codons (Lipke and Kurjan, 1992).

Sexual agglutination has been studied in several other species of budding yeast including *Hansenula wingei, Pichia amethionina* and *Saccharomyces kluyveri*. The structural features of the agglutinins, peptide pheromones and pheromone receptors have been conserved across the different yeast species, although the agglutinin interactions themselves are highly species-specific (Lipke and Kurjan, 1992).

1.7 TARGETING THE CELL WALL IN *S. cerevisiae*

Display of heterologous proteins on the surface of *S. cerevisiae* is well suited for eukaryotic cell surface proteins that require endoplasmic reticulum specific post-translational processing such as glycosylation and disulphide bond formation (Boder and Wittrup, 1997). Eukaryotic expression systems are often able to overcome the unpredictable biases observed when expressing eukaryotic proteins in prokaryotic expression systems (Knappik and Plückthun, 1995). The ability to display eukaryotic proteins on the surface of *Saccharomyces cerevisiae* has created new development opportunities such as immobilisation of enzymes (Klis *et al.*, 1994; Murai *et al.*, 1997a; Murai *et al.*, 1997b), screening of polypeptide libraries (Boder and Wittrup, 1997) and the creation of live vaccines (Schreuder *et al.*, 1996).

The first heterologous protein to be targeted to the cell wall of *S. cerevisiae* was the enzyme α-galactosidase from *Cyamoposis tetragoroloba* (guar) seeds (Schreuder *et al.*, 1993). This construct consisted of an invertase signal sequence, the α-galactosidase gene and the C-terminal half of the α-agglutinin (Figure 1.10).

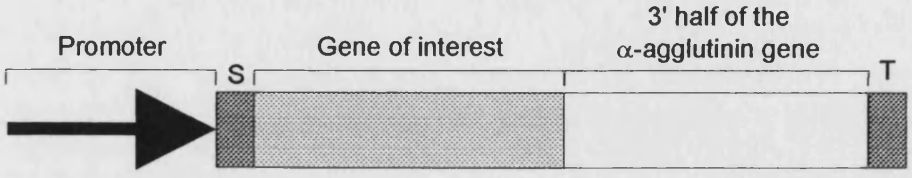


Figure 1.10 - The general DNA construct for expression of a protein on the surface of *S. cerevisiae*.

All DNA constructs used to express various proteins on the surface of *S. cerevisiae* (Table 1.2) contain a promoter, a secretion signal sequence (S), the gene of the protein of interest, the 3' anchorage domain of the α-agglutinin gene and a terminator sequence (T).

This fusion was expressed onto the surface of *S. cerevisae* under the control of a constitutive phosphoglycerokinase (PGK) promoter. Analysis of the cell wall-extracted material showed that 70% of the α-galactosidase/α-agglutinin protein fusion was covalently attached to the β-1,6-glucan component of the cell wall. Furthermore, immunofluorescence experiments confirmed the α-galactosidase protein was on the surface of *S. cerevisiae.* However, the detectable amount of expression varied greatly between different cells in the population due to changes in the copy number of the expression vector (Schreuder *et al.*, 1996).

To date, seven proteins have been displayed on the surface of *S. cerevisiae* using the C-terminal half of the α-agglutinin gene (Table 1.2). In all cases, the N-terminal half of the constructs were detectable immunologically on the surface of the intact cells and the fusions were tightly linked to the β-glucan component of the cell wall.

Table 1.2 - Proteins displayed on the surface of *S. cerevisiae* using the C-terminal half of the α-agglutinin gene.

N-terminal gene	Reference
α-galactosidase	Schreuder *et al.*, 1993
VH antibody fragment	Hamers-Casterman *et al.*, 1993
Lipase	Klis *et al.*, 1994
Cutinase	Klis *et al.*, 1994
Hepatitis B surface antigen	Schreuder *et al.*, 1996
Cellulase	Murai *et al.*, 1997a
Glucoamylase	Murai *et al.*, 1997b

Recently, the a-agglutinin gene has also been used to target the cell surface of *S. cerevisiae* (Boder and Wittrup, 1997; Kieke *et al.*, 1997). In these cases, the heterologous protein was fused to the C-terminus of the a-agglutinin binding domain (Aga2p). In 1997, Boder and Wittrup displayed a mutated antibody fragment library on the yeast cell surface using this method and high affinity antibodies were selected directly using fluorescence-activated cell sorting (FACS). This a-agglutinin surface display system has also been used to express high levels of a T-cell receptor on the surface of yeast (Cho *et al.*, 1998). These engineered yeast cells were shown to act as 'pseudo' antigen-presenting cells and induce T-cell activation.

Additionally, other cell wall proteins such as flocculin gene (*FLO1*) have been shown to be effective in cell surface display, although large differences in immobilisation efficiency compared to the C-terminal half of the α-agglutinin were observed (van der Vaart *et al.*, 1997).

1.8 RATIONALE OF THE PROJECT

Antibodies intended for diagnostic and therapeutic purposes must bind tightly and rapidly to the relevant antigen to avoid undesirable side reactions. High affinity monoclonal human antibodies demonstrate great therapeutic potential in the treatment of diseases such as cancer (Schlom *et al.*, 1992). The fundamental reason behind the conception of this project was to eventually create an antibody improvement scheme that could generate high affinity monoclonal human antibodies faster and more efficiently than any other improvement system already established.

The improvement of antibodies *in vivo* occurs in the germinal centre by a process known as affinity maturation (Section 1.3.3). The germline antibody DNA is mutated inside a B-cell and then antibodies containing point mutations are expressed on the B-cell's surface. B-cells displaying advantageously mutated antibodies are selected due to their ability to interact with antigen-presenting follicular dendritic cells. Positively-selected B-cells often undergo further rounds of mutation and selection to generate

the high affinity antibodies which are characteristic of a secondary immune response. Ultimately, this project aims to produce an antibody improvement system based on, and as effective as, the germinal centre reaction in the human body.

S. cerevisiae is the chosen organism for antibody expression in this project (Section 1.5). The main reasons for selecting the yeast *S. cerevisiae* are firstly its genetics are very well established and therefore its genome can be easily manipulated. The ability to manipulate the yeast genome was vital for success in the later stages of this project. Secondly unlike phage, single yeast cells can be isolated using fluorescence-activated cell sorting (FACS) which was an essential requirement in the design of Part I of this project. Finally, *S. cerevisiae* is eukaryotic and therefore is able to undertake post-translational modifications such as disulphide bond formation, which is extremely important for antibody production.

To mimic the antibody-displaying properties of B-cells, the C-terminal anchorage domain of the α-agglutinin gene was used to show that antibody fragments can be functionally expressed on the cell surface of the yeast, *S. cerevisiae*. Unfortunately, the yeast FACS system devised to imitate the positive-selection of B-cells in the germinal centre reaction was never demonstrated in this work. This was due to the fact that an *S. cerevisiae* FACS-selection system had already been successfully established (Boder and Wittrup, 1997). Therefore, the second part of this project was designed to create a new antibody selection system based on the mating characteristics of yeast agglutinin mutants (Section 1.6.2).

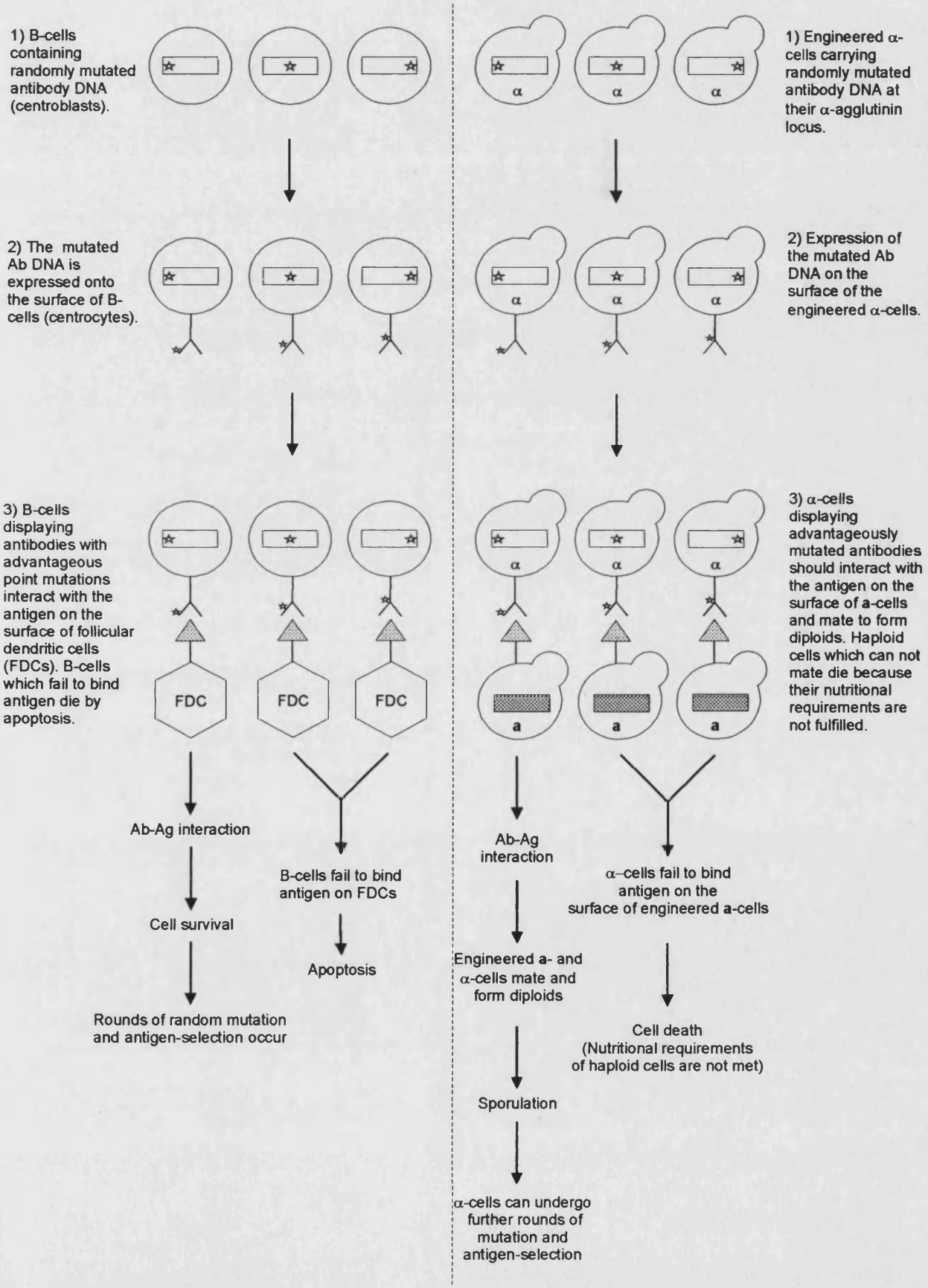
It was intended to mimic the germinal centre reaction by using the α-agglutinin gene to display antibodies on the surface of α-cells and the **a**-agglutinin gene to create the antigen-presenting system on the surface of **a**-cells (Figure 1.11). This would be done by integrating the antibody and antigen DNA at the relevant agglutinin loci in the appropriate haploid cell genomes (Schreuder *et al.*, 1996) to generate surface-display systems that have constant expression levels throughout the cell populations. Mating of these engineered haploid strains will hopefully create a selection system that is able to choose cells which are expressing high affinity antibodies and will

not be affected by cells which are displaying moderate affinity antibodies at higher expression levels. In order to efficiently integrate a mutant antibody library into the α-agglutinin locus a novel site-specific recombination system was developed. A large surface display library ought to improve the chances of selecting for high affinity antibodies directly (Perelson and Oster, 1979; Griffiths *et al.*, 1994).

The proposed selection process in Part II of this project will depend on the antigen-antibody interaction between the haploid mating-type display systems. An antibody-antigen interaction was chosen which has a similar affinity constant to the **a**-/α-agglutinin interaction (Ka = 10^9M^{-1}) in order to reproduce the agglutination process as closely as possible. Engineered α-cells displaying low affinity antibodies should not be able to interact sufficiently to produce cell adhesion and die because their nutritional requirements will not be fulfilled on subsequent selection on nutritionally-deficient medium. The positively-selected diploids cells will then undergo meiosis (in nutritional starvation conditions) and α-cells expressing high affinity antibodies will be isolated to undergo further rounds of mutation and antigen-selection. Eventually, it was anticipated that an 'in-cell' mutational system would be developed to avoid the extensive DNA manipulations, which are required between antigen-selection steps. Producing an antibody improvement system that can occur totally within yeast cells would be a very attractive prospect.

Ideally this yeast 'germinal centre reaction' will produce high affinity human antibodies which can be used directly for therapeutic applications, thus by-passing the requirement for immunisation (Section 1.4.1) and humanisation (Section 1.4.2).

Figure 1.11 A diagrammatic comparison between the antibody improvement system *in vivo* (left) and the proposed artificial yeast germinal centre strategy (right). The stars represent the point mutations which could occur throughout the germline antibody DNA.



1.9 AIMS OF THE PROJECT

The overall aim of this project is to demonstrate an 'in-cell' affinity maturation strategy using the yeast, *S. cerevisiae*. Ultimately, this strategy will mimic the human germinal centre reaction (Section 1.3.3) as closely as possible to produce high affinity human antibodies for therapy. If successful, this process will by-pass the requirement for immunisation (Section 1.4.1) and become the eukaryotic equivalent to phage display technology (Section 1.4.4).

The first aim of this project is to show that antibody fragments (scFv and Fab) can be expressed on the surface of *S. cerevisiae* and are accessible to their targets (Part I). It is proposed that the C-terminal half of the α-agglutinin gene can be used for the desired antibody surface display system by covalent attachment of the antibody fragments to the β-glucan component of the cell wall (Section 1.6.2). This technology has already been used to display a number of heterologous proteins on the yeast cell surface (Section 1.7). The antibody fragment expression vectors will consist of the α-agglutinin leader sequence which will target the antibody fragments to the secretory pathway, the C-terminal half of the α-agglutinin gene (except in the case of the Fab light chain), a selectable marker and a strong yeast inducible promoter (Chapter 3). The antibody fragments will have affinity for the fluorescent antigen, Fluorescein-5-isothiocyanate (FITC), therefore functionality will be shown directly using immunofluorescence (Chapter 4). Part I of this project was designed with the intention of developing a fluorescence-activated cell sorting (FACS) selection system that could isolate yeast cells expressing high affinity antibodies from a mixed population. Unfortunately, this objective was never achieved due to the fact a FACS-based yeast selection system had already been established (Boder and Wittrup, 1997). Therefore, although the overall aim of this project remained the same, the goal of the second half of this project will be to develop an alternative selection strategy that utilizes the mating attributes of agglutinin mutants.

Hence, the second aim of this project is to show that the germinal centre reaction can be mimicked using opposite mating-type haploid cells has antigen and antibody surface display systems (Figure 1.11; Part II). For this selection process to succeed it was decided constant surface expression of the antibody (and antigen) agglutinin fusions was required to avoid problems distinguishing between cells expressing high affinity antibodies from cells displaying moderate affinity antibodies at higher levels. To achieve constant surface expression, the objective will be to integrate the antibody and antigen genes at the appropriate positions in the agglutinin loci. The Cre-*lox* site-specific recombination system will be recruited to integrate the antibody DNA into the α-agglutinin locus at an efficiency required for mutant library construction (Chapter 5).

The antigen-antibody system used will be the well-characterised scFv D1.3-hen egg white lysozyme interaction. The **a**-/α-agglutinin and scFv D1.3/lysozyme interactions have similar affinity constants (Ka ~ $10^9 M^{-1}$), therefore the germinal centre reaction will be copied by attempting to mate engineered **a**- and α-cells displaying the lysozyme and scFv D1.3, respectively (Figure 1.11). Engineered α-cells displaying low affinity antibodies ought not to bind the antigen presented on **a**-cells and die because their nutritional requirements will not be met. Only mated (diploid) cells will be able to grow on the selective media used (Chapter 6). This diploid selection strategy is intended to mimic the antigen-selection process in the germinal centre reaction. Additionally, once this selective mating process is established, the future direction of this project will be to develop an 'in-cell' mutational switch to avoid all the extensive DNA manipulations needed between antigen-selection steps. Ultimately, rounds of mutation and antigen-selection within yeast cells should lead to direct isolation of high affinity human antibodies.

CHAPTER 2

GENERAL MATERIALS AND METHODS

This chapter details the general experimental techniques used throughout this project. Any materials and methods specific to a particular chapter are described within that chapter.

2.1 MATERIALS

2.1.1 Chemical reagents

Chemical reagents (analytical grade) were obtained from Sigma, Fisons Ltd. or BDH Ltd. except when stated otherwise. Water used throughout this project was Millipore Q-Plus filtered water.

2.1.2 Molecular biological enzymes and reagents

All restriction endonucleases, $Vent_R$ DNA polymerase, T4 DNA ligase, T4 polynucleotide kinase and β-agarase I were purchased from New England Biolabs Inc. *Taq* DNA polymerase was obtained as BIOTAQ™ polymerase from Bioline Ltd and *Taq*/*Pwo* DNA polymerase mix was obtained as Expand™ High Fidelity PCR system from Boehringer Mannheim. Ultrapure 2'- deoxynucleoside 5'-triphosphates (dNTPs) were purchased from Pharmacia Biotech. Ultrapure™ agarose, Ultrapure™ low melting point agarose, 100 b.p. and 1 k.b.p. DNA ladders were supplied by Life Technologies. Mineral oil, salmon sperm DNA and pancreatic RNase were purchased from Sigma.

2.1.3 Culture media and antibiotics

Bacto-yeast extract, Bacto-peptone, Bacto-tryptone, and Bacto-yeast nitrogen base without amino acids were all purchased from Difco Laboratories. Bacto-agar was either purchased from Difco Laboratories or BDH Ltd. All amino acids were obtained from Sigma and wherever applicable, HCl salts of amino acids were ordered. Glucose, raffinose, galactose (minimum 99% pure) and the antibiotics, ampicillin and geneticin (G418) were also purchased from Sigma.

2.1.4 Other materials

Nitrocellulose membrane was obtained as Hybond™-C from Amersham International and the 3MM paper from Whatman Scientific Ltd. The standard protein markers were either purchased from Amersham International (High and low molecular weight Rainbow™ protein markers) or New England Biolabs Inc. (Broad range prestained protein marker). Protogel™ was supplied by National Diagnostics Ltd. (New Jersey, USA) and was a premade 37:5:1 acrylamide:bisacrylamide solution, supplied at a 30% (w/v) acrylamide concentration. The 96-well ELISA plates and lids were obtained as Falcon 3072 tissue culture plates from Becton Dickinson. Cycloheximide and 3, 3', 5, 5'-tetramethyl benzidine (TMB) were purchased from Sigma.

2.2 MOLECULAR BIOLOGICAL METHODS

2.2.1 Bacterial strains and culture conditions

Escherichia coli strain TOP10F' (F'{lacI^q *Tn*10(Tet^R)} *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) ϕ80*lacZ*ΔM15 Δ*lacX*74 *deo*R *rec*AI *ara*D139Δ(*ara-leu*)7697 *gal*U *gal*K *rsp*L *end*A1 *nup*G) was the standard strain used throughout this project and obtained as part of the Eukaryotic TA cloning kit (Version 1.0) from Invitrogen.

Plasmids, which were supplied in other strains, are described in the appropriate chapters.

E. coli strains were cultured at 37°C in Luria broth [(LB) 10g/l Bacto-tryptone, 5g/l Bacto-yeast extract, 10g/l sodium chloride] liquid media [250

revolutions per minute (r.p.m.)], or on LB-agar plates (supplemented with 15g/l Bacto-agar) overnight.

For *E. coli* strains carrying plasmids conferring ampicillin resistance, the medium was supplemented with ampicillin to a final concentration of 100μg/ml.

2.2.2 *Saccharomyces cerevisiae* strains and culture conditions

All *Saccharomyces cerevisiae* strains are described in the appropriate chapters and were cultured at 30°C in Yeast-extract peptone dextrose [(YPD) 10g/l Bacto-yeast extract, 20g/l Bacto-peptone, 20g/l glucose] liquid media (250r.p.m.) overnight, or on YPD-agar plates (supplemented with 15g/l Bacto-agar) for 24-48 hours.

Plasmids carrying the *HIS3*, *LEU2*, or *URA3* selectable markers were transformed into *S. cerevisiae* strains with the corresponding auxotrophic mutations. Positive transformants were cultured in Synthetic Complete (SC) liquid media without the relevant amino acids (6.7g/l Bacto-yeast nitrogen base without amino acids, 20g/l glucose, 2g/l "drop-out" mix) at 30°C for approximately 24 hours. The "drop-out" mix was a combination of the following ingredients without the amino acids* which complemented the relevant plasmid:

Adenine	0.5g	Leucine*	10.0g
Alanine	2.0g	Lysine	2.0g
Arginine	2.0g	Methionine	2.0g
Asparagine	2.0g	Phenylalanine	2.0g
Aspartic acid	2.0g	Proline	2.0g
Cysteine	2.0g	Serine	2.0g
Glutamine	2.0g	Threonine	2.0g
Glutamic acid	2.0g	Tryptophan	2.0g
Glycine	2.0g	Tyrosine	2.0g
Histidine*	2.0g	Uracil*	2.0g
Inositol	2.0g	Valine	2.0g
Isoleucine	2.0g		

The *S. cerevisiae* strains were also grown on SC lacking amino acids* plates (supplemented with 20g/l Bacto-agar) at 30°C for 4 days. The amino acids absent in the SC plates were produced by the appropriate transformed plasmids and were essential for the nutritional requirements of the chosen yeast strain.

Any other culture conditions/media used are described in the relevant chapters.

2.2.3 Glycerol stocks

All *E. coli*/*S. cerevisiae* strains were stored at -70°C in 16% (v/v) glycerol.

2.2.4 Design of oligonucleotides

Oligonucleotide primers were designed to bind to the designated DNA template with the following factors taken into consideration (McPherson *et al.*, 1992):

i) The primers consisted of at least 18 bases of homology to the target DNA to ensure good hybridisation. If a choice was available in the exact DNA sequence, the region of the greatest base diversity was chosen.

ii) The primer (G+C) base content was at least 50% to ensure stability of the primer-template duplex.

iii) Their annealing temperatures were calculated using the Wallace equation: Tm = 4(G+C) + 2(A+T) [Applied Biosystems manual - Automated DNA sequencing chemistry guide] and was always above 45°C. If possible, a primer pair with similar annealing temperatures was used.

iv) The primer pairs were checked for regions of complementarity to each other to avoid 'primer-dimer' formation. Primer-dimer formation can lead to a reduction of the amount of available primers within a reaction.

v) Individual primers were also checked for self-complementarity to prevent primers annealing to themselves.

If restriction sites were required at the ends of the PCR products, 6-10 extra nucleotides were added after the enzyme recognition sites to increase the efficiency of cutting. The reference section in the latest New England Biolabs catalogue ('Cleavage Close to the End of DNA Fragments') was

consulted to decide the number and sequence of bases to flank the recognition sites.

2.2.5 Synthesis of oligonucleotides

Oligonucleotides used in Chapters 3/4 were synthesized by Severn Biotech. Ltd. and were supplied HPLC purified as a lyophilised pellet. The pellet was reconstituted in distilled water to a final concentration of 1 nmol/μl, unless otherwise stated and stored at -20°C.

Oligonucleotides used in Chapters 5/6 were synthesized by Perkin-Elmer Applied Biosystems and were supplied in 20% (v/v) acetonitrile/water and stored at -20°C.

An aliquot of each primer was used throughout as a working stock to prevent contamination and repeated freeze-thawing.

2.2.6 Quantitation of oligonucleotides

The concentration of each oligonucleotide was confirmed by measuring the absorbance at 260nm on a Cecil Ce6600 Multimode Computis UV spectrophotometer using a matched pair of Hellma quartz curvettes (pathlength 10mm). The primer concentration was calculated using the following formula (provided by Perkin-Elmer):

$$\text{Primer concentration} = \text{pmol}/\mu\text{l} = \frac{100 \times A_{260}}{(1.54nA+0.75nC+1.17nG+0.92nT)}$$

(n = number of A's, C's, G,'s or T's)

2.2.7 The Polymerase Chain Reaction (PCR)

The precise cycling parameters for polymerase chain reactions (PCRs) were optimised and are listed in the appropriate sections, but the following basic guidelines were always applied:

i) DNA denaturation was carried out at 94°C.

ii) The annealing temperature was calculated for the primers using the Wallace rule (Section 2.2.4). If a PCR product was not amplified, the annealing temperature was decreased, and if any non-specific products were observed, the annealing temperature was increased.

iii) For amplification to 3k.b.p., an elongation temperature of 72°C was used. For any amplification above 3k.b.p., the elongation temperature was lower to 68°C.

iv) The elongation time depended on the PCR product length and the following table was adhere to:

Elongation Time	45sec	1min	2min	4min	8min
PCR fragment length (k.b.p)	>0.75	1.5	3	6	10

v) The number of cycles were kept to a minimal to reduce the opportunity for errors.

Note: These guidelines are based on the Expand™ High Fidelity PCR System information sheet supplied by Boehringer Mannheim.

PCRs were performed on a M.J. Research Inc. PTC-100 programmable thermal controller using either BIOTAQ™ *Taq*, $Vent_R$ or Expand™ DNA polymerases. Both $Vent_R$ and Expand™ DNA polymerases exhibit 3'-5' exonuclease proofreading activity, therefore increasing the fidelity of DNA synthesis compared to *Taq* DNA polymerase. The following table shows the error rates of the various DNA polymerases [figures were obtained from the Boehringer Mannheim Expand ™ High Fidelity PCR System data sheet and the New England Biolabs catalogue 1998/99 (page 76)]:

DNA polymerase	$Vent_R$	Expand™	*Taq*
Error rate (10^{-6} errors/b.p./duplication)	57	8.5	285

A) *$Vent_R$ polymerase chain reactions*

For a $Vent_R$ DNA polymerase reaction the conditions were: 100pmol of each primer, 5µl of 10x NEB $Vent_R$ buffer [100mM KCl, 200mM Tris-HCl (pH8.8),

100mM $(NH_4)_2SO_4$, 0.1% (v/v) Triton X-100], 5µl of 10x stock dNTPs (300µM dATP, 300µM dCTP, 300µM dGTP, 300µM dTTP) and 0.5µl of bovine serum albumin (10mg/ml). These reagents were mixed in a 0.2ml thin-walled tube and template DNA (50-250ng) was added. Finally, the mixture was made up to a volume of 49.5µl with water and 0.5µl of $Vent_R$ polymerase (2units/µl) was added.

B) *Taq polymerase chain reactions*

For a *Taq* DNA polymerase reaction the conditions were: 25pmol of each primer and 10µl of 5x *Taq* buffer [80mM $(NH_4)_2SO_4$, 335mM Tris-HCl (pH8.8), 0.05% (v/v) Tween 20, 100µM dATP, 100µM dCTP, 100µM dGTP, 100µM dTTP, 7.5mM $MgCl_2$]. These reagents were mixed in a 0.2ml thin-walled tube and 50-250ng of template DNA was added. Finally, the mixture was made up to a volume of 49.5µl with water and 0.5µl of BIOLINE™ *Taq* polymerase (5 units/µl) was added.

C) *Expand™ polymerase chain reactions*

For an Expand™ DNA polymerase reaction the conditions were: 25pmol of each primer, 5µl of 10x Expand buffer [20mM Tris-HCl (pH7.5), 100mM KCl, 1mM dithiothreitol, 0.1mM EDTA, 0.5% (v/v) Tween 20, 0.5% (v/v) Nonidet P40, 15mM $MgCl_2$] and 5µl of 10x stock dNTPs (200µM dATP, 200µM dCTP, 200µM dGTP, 200µM dTTP). These reagents were mixed in a 0.2ml thin-walled tube and 50-250ng of template DNA was added. Finally, the mixture was made up to a volume of 49.25µl with water and 0.75µl of Expand™ DNA polymerase (3.5 units/µl) was added.

All PCRs were run with a negative control (no DNA template) to check for contamination. If available a positive control on an existing template was also carried out to check various parameters such as buffers, enzymes and temperature cycles. If the programmable thermal controller without the insulated lid was used, the PCRs were overlaid with a drop of mineral oil to prevent evaporation. All PCRs were kept at 4°C until analysis and each PCR (5µl) was examined using agarose gel electrophoresis (Section 2.2.8),

unless otherwise stated. If the PCR contained no non-specific bands when analysed, it was purified directly using the Wizard™ PCR Purification System (Promega) according to the manufacturer's instructions.

2.2.8 Agarose Gel Electrophoresis

DNA fragments were visualised using agarose gel electrophoresis. The agarose concentration was adjusted from 0.8-3.0% (w/v) to resolve a wide range of DNA molecules (10k.b.p. - 50b.p., respectively). The DNA fragments were stained using ethidium bromide and visualised directly using long or short wavelength ultraviolet (UV) light.

Gels were prepared by dissolving the appropriate amount of Ultrapure™ agarose in 50ml of TBE buffer (90mM Tris-borate, 2mM EDTA) by boiling in a microwave oven. Ethidium bromide was added to a final concentration of 0.5μg/ml when the agarose had cooled to 50-60°C. The agarose was then poured into a Flowgen (10cm x 7cm) mini gel unit, which already contained spacers and a well-forming comb. Once set the comb and spacers were removed and the gel was submerged in TBE. The DNA samples were mixed with 6x loading buffer [0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol, 40% (w/v) sucrose] and loaded onto the gel. Appropriate DNA markers were loaded either side of the samples and a constant voltage of 80V (8.0V/cm) was applied. The gel was run until the bromophenol blue had migrated sufficiently and then visualised with short wavelength UV light.

To purify a DNA fragment from an agarose gel, low-melting point (LMP) agarose was used. LMP agarose gels were prepared by dissolving the relevant amount of LMP agarose in 75ml of TAE buffer (40mM Tris-acetate, 1mM EDTA) by heating in a microwave oven. Once cooled to 50-60°C, ethidium bromide was added to a final concentration of 0.5μg/ml and poured in the cold room (4°C). The solidified agarose was submerged in TAE buffer and the DNA samples were mixed with 6x LMP loading buffer [0.25% (w/v) bromophenol blue, 40% (w/v) sucrose]. The gel was ran at 40V (4V/cm) in the cold room (4°C), until the bromophenol blue had migrated sufficiently. The DNA fragments were briefly visualised with a long wavelength hand-held UV lamp to avoid DNA damage.

2.2.9 Recovery of DNA from LMP agarose gels

Two techniques were used for purifying DNA from LMP agarose gels:

1) β-agarase digestion.
2) DNA extraction kits.

The relevant DNA fragment was cut out of a LMP agarose gel with a clean razor blade and melted at 80°C for 5 minutes.

If a small quantity of DNA was available or the DNA fragment was going to be sequenced, β-agarase was used.

A) *β-agarase digestion of LMP agarose*

β-agarase buffer [10x (100mM Tris-HCl (pH6.5), 10mM EDTA)] was added to the melted LMP agarose to a final concentration of 10% (w/v). This mixture was reheated for 5 minutes, mixed and then allowed to cool to 40°C. Two units of β-agarase were added to digest 200μl of 1% LMP agarose gel and incubated at 40°C, overnight. Sodium acetate (pH 5.2) was added to a final concentration of 0.3M and the mixture was incubated on ice for 15 minutes. The undigested agarose was removed by centrifugation at 13000r.p.m. for 15 minutes in an MSE Micro-Centaur microfuge and the supernatant was transferred to a clean 1.5ml Eppendorf tube. The DNA was precipitated with two volumes of ice-cold isopropanol, mixed thoroughly and incubated at -20°C for 30 minutes. The supernatant was removed by re-centrifugation and the pellet was washed with 1ml of 70% ethanol. The supernatant was then carefully aspirated off and the pellet was dried at room temperature. The DNA pellet was then resuspended in the appropriate amount of water for subsequent manipulations.

B) *DNA extraction kits*

If the DNA fragment was above 3k.b.p., the melted gel was purified using either the Qiaquick™ gel extraction kit (Qiagen) or the Wizard™ DNA Clean Up System (Promega) according the manufacturer's instructions.

If the DNA fragment was less than 3k.b.p., the Wizard™ PCR Purification System (Promega) was used by following the supplied instructions.

2.2.10 Alcohol precipitation of DNA

A 3M sodium acetate (pH 5.2) solution was added to the DNA to a final concentration of 300mM. Three volumes of absolute ethanol were added, the solution was vortexed and then incubated for one hour or overnight at -20°C. The precipitated DNA was collected at 13000r.p.m. for 5 minutes at 4°C, using an MSE Micro-Centaur microfuge. The supernatant was removed and the DNA pellet was washed once in 1ml of 70% ethanol. The DNA pellet was then air-dried, resuspended in the required amount of water and stored at -20°C.

2.2.11 Quantitation of DNA

A DNA solution was diluted with a suitable amount of water and the absorbance was measured at 260nm on Cecil Ce6600 Multimode Computis UV spectrophotometer. One absorbance unit at 260nm is equivalent to 50μg/ml of double stranded DNA (Sambrook *et al.*, 1989).

In practice, the DNA concentration of samples was often estimated by comparison to a known concentration of DNA marker using agarose gel electrophoresis (Section 2.2.8).

2.2.12 Restriction endonuclease digestion of DNA

All restriction enzymes were obtained from New England Biolabs (NEB) and used according to the manufacturer's instructions. If two enzymes had a compatible buffer, a double digest was performed in one step. A reaction mixture consisted of 5μl of the required 10x NEB buffer and 1μl of each of the restriction enzymes (3-20 units). Some NEB enzymes required the addition of bovine serum albumin to a concentration of 100μg/ml. The reaction mixtures were incubated at the required temperature for 16 hours and then heat inactivated at 65°C for 20 minutes. Digestions were performed using 1-1.5μg of DNA in a final reaction volume of 50μl, unless otherwise stated.

2.2.13 Ligations of sticky ended DNA

Ligation mixtures were prepared using 6 Weiss units (400NEBunits/μl) of T4 DNA ligase, 1μl of 10x NEB ligase buffer [500mM Tris-HCl (pH7.5), 100mM $MgCl_2$, 100mM dithiothreitol, 10mM ATP, 250μg/ml bovine serum albumin] and insert and vector DNA with compatible cohesive-ends, into a final reaction volume of 10μl. Normally the molar ratio of insert to vector was 3 to 1. To assess the degree of vector religation, a cut vector-only control was prepared each time. These mixtures were incubated at 16°C for 16 hours and then heat-inactivated at 65°C for 10 minutes. Half of each ligation mixture (5μl) was directly transformed into competent TOP10F' *E. coli* cells and the remaining amount was stored at -20°C.

2.2.14 Preparation of competent *E. coli* cells (Sambrook *et al.*, 1989)

TOP10F' *E. coli* cells (50μl) from a glycerol stock were plated onto a LB agar plate and incubated overnight at 37°C. A loopful of these cells was used to inoculate 3ml of LB medium and the resultant cell suspension was grown to saturation overnight at 37°C. A 500μl volume of this overnight culture was used to inoculate 50ml of LB medium and was grown at 37°C, until an absorbance (at 450nm) of 0.4 units was reached. The cells were then centrifuged at 3000r.p.m. (1500g) for 5 minutes using a GLC4 General Laboratory Centrifuge (Sorvall Instruments) and the pellet was resuspended in 25ml of ice-cold 50mM calcium chloride. This suspension was left on ice for 20 minutes and then re-centrifuged. The cell pellet was resuspended in 5ml of 50mM calcium chloride/18% (v/v) glycerol, snap-frozen in 100μl aliquots and stored at -70°C.

2.2.15 Transformation of foreign DNA into competent *E. coli* cells

A ligation mixture (5μl) or plasmid DNA (1μl) was added to a 100μl aliquot of competent *E. coli* cells, which had been gently thawed on ice. The cells were incubated on ice for 1 hour, heatshocked at 42°C for 90 seconds and then placed on ice again for 2 minutes. A 400μl volume of SOC media (20g/l Bacto-tryptone, 5g/l Bacto-yeast extract, 0.5g/l NaCl, 2.5mM KCl, 10mM

$MgCl_2$, 20mM glucose) was added and the cells were left to shake (250 r.p.m.) at 37°C for 1 hour. An aliquot of each transformation (200µl) was then plated onto an LB plate, supplemented with 100µg/ml ampicillin and allowed to dry. The plates were then inverted and incubated at 37°C, overnight (this method was adapted from Eukaryotic TA cloning kit instructions, Invitrogen).

2.2.16 PCR screening of positive transformants

In order to check that the transformants contained the correct plasmid or integrated construct, PCR screening was undertaken (McPherson *et al.*, 1992).

A numbered *E. coli/S. cerevisiae* colony was selected using a sterile toothpick and then transferred into a PCR tube containing the relevant *Taq* polymerase PCR reaction [Section 2.2.7 (20µl)]. The toothpick was left in the PCR tube for 2 minutes and then transferred to a numbered culture tube containing the appropriate selective media.

The PCR cycling parameters were: [94°C, 5min; (94°C, 1min; 50°C, 1min; 72°C, 1min) x 25; 72°C, 5min] unless otherwise stated and 5µl of each of the PCR products were analysed by agarose gel electrophoresis (Section 2.2.8).

2.2.17 Small scale preparation of plasmid DNA ('Minipreps')

Small quantities of plasmid DNA were prepared from an overnight culture of *E. coli* cells carrying the plasmid of interest, using the Wizard™ miniprep DNA purification system (Promega), according to the manufacturer's instructions.

When plasmid DNA was required for DNA sequencing, the S.N.A.P miniprep kit (Invitrogen) was used and the manufacturer's instructions were followed exactly.

2.2.18 Automated DNA sequencing and sequence analysis

For automated sequencing, 250-500ng of plasmid DNA or linear DNA was required. A sequencing primer (3.2pmol) was added to a 0.2ml thin-walled tube containing the DNA template and a total volume of 6µl was completed

with water. The samples were sequenced on an ABI PRISM 377 DNA Sequencer, using the fluorescent dye terminator protocol that was based upon the chain-terminator sequencing method (Sanger *et al.*, 1977).

The resulting DNA sequence chromatographs were analysed and checked against the published DNA sequence for any base mutations, using the Sequence Editor (SeqEd™) program version 1.03 (Perkin-Elmer, Applied Biosystems Division).

2.2.19 Preparation of pancreatic RNase A

Pancreatic RNase A (100mg) was dissolved in 10ml of 10mM Tris-HCl (pH7.5), 15mM NaCl and heated at 100°C for 15 minutes. The mixture was cooled to room temperature, dispensed into 0.5ml aliquots and stored at -20°C.

2.2.20 Preparation of single stranded carrier DNA (Johnston, 1994)

Salmon sperm DNA (0.1g) was dissolved into 10ml TE buffer [10mM Tris-HCl (pH8.0), 1mM EDTA] and sonicated, using a 19mm titanium probe, for 30 seconds at 300-500W (MSE Isowatt Ultrasonic Disintegrator, Fisons). Once sonicated sufficiently, the DNA was extracted with phenol:chloroform:isoamyl alcohol (25:24:1) and centrifuged at 2000g for 5 minutes, using a Megafuge 1.0 centrifuge (Heraeus instruments). The top aqueous layer was removed and transferred to a clean 50ml-centrifuge tube. The DNA was precipitated by adding 1ml 3M sodium acetate (pH 5.2) and 25ml ice-cold 100% ethanol. This solution was vortexed and centrifuged at 5000g for 10 minutes. The DNA pellet was washed twice with 25ml 70% ethanol and allowed to air-dry. The pellet was resuspended in 10ml TE buffer (10mg/ml DNA solution), aliquoted and stored at -20°C.

2.2.21 Preparation of yeast genomic DNA (Adams *et al.*, 1996)

A 10ml culture of the relevant strain of *Saccharomyces cerevisiae* was grown to saturation in YPD at 30°C, overnight. The cells were collected for 2 minutes at 3000r.p.m. (1500g) using a Sorvall GLC-4 general laboratory centrifuge. The supernatant was removed and the cells were resuspended in

0.5 ml of water. The cell suspension was transferred to a 1.5ml screw-capped Eppendorf tube and centrifuged at 13000r.p.m. for 30 seconds, using a MSE Micro-Centaur microfuge. The supernatant was removed and the pellet was resuspended in the residual liquid. Cell resuspension buffer [0.2ml; 2% (w/v) Triton X-100, 1% (w/v) SDS, 100mM NaCl, 10mM Tris-HCl (pH8.0)], 0.2ml of phenol:chloroform:isoamyl alcohol (25:24:1) and 0.3g of acid-washed glass beads (0.2 microns) were added and vortexed for 4 minutes. Water (0.2ml) was added, the mixture was vortexed briefly and centrifuged for 5 minutes at 13000r.p.m. The top layer was transferred to a new 1.5ml Eppendorf tube and 1ml of 100% ethanol was added to precipitate the DNA. The DNA solution was then vortexed and centrifuged for 2 minutes at 13000r.p.m. The supernatant was discarded and the DNA pellet was resuspended in 0.4ml of water. Five microlitres of a 10mg/ml RNase A solution was added and incubated at 37°C for 30 minutes. Approximately 1ml of 100% ethanol and 10µl of 4M ammonium acetate were added, mixed and centrifuged for 2 minutes at 13000r.p.m. The DNA pellet was washed with 70% ethanol and either air-dried, or dried under vacuum using a Stratech Scientific SpeedVac concentrator. The DNA pellet was then resuspended in the required volume of water for subsequent manipulations.

2.2.22 Transformation of *S. cerevisiae* cells with plasmid DNA (Johnston, 1994)

The relevant strain of *S. cerevisiae* was grown on a YPD plate at 30°C, overnight. Approximately 50µl of cells were scraped off the YPD plate and resuspended in 1ml of water. The cell suspension was then centrifuged for 1 minute at 13000r.p.m. and the supernatant was discarded. The cell pellet was washed with 1ml of water and re-centrifuged. The pellet was then washed with 0.5ml of 100mM lithium acetate, re-centrifuged and then resuspended in an equal volume of 100mM lithium acetate. Five microlitres of freshly boiled single-stranded DNA (10mg/ml) and up to 5µg of plasmid DNA was added to 50µl of lithium-acetate treated cells and vortexed. A 40% (w/v) polyethylene glycol 3350/100mM lithium acetate solution (300µl) was

added and vortexed again. The mixture was then heatshocked at 42°C for 20 minutes, centrifuged for 10 seconds at 13000r.p.m. and the cell pellet was resuspended in 1ml of water. Aliquots of the transformed cells (250µl) were then plated onto the appropriate Selective Complete medium.

2.2.23 Super efficient transformation of *S. cerevisiae* cells with plasmid DNA (Adams *et al.*, 1996)

A single colony of the appropriate strain of *S. cerevisiae* was used to inoculate 5ml of YPD and grown overnight at 30°C. The cell density of the pre-culture was estimated using a haemocytometer and diluted to 5 x 10^6 cells/ml in 100ml of YPD. The cells were grown for two additional cell divisions (~3 hours) and collected by centrifugation at 3000r.p.m. (1500g) for 5 minutes using a GLC4 general laboratory centrifuge. The cells were washed twice with 10ml of water and once with 10ml of 100mM lithium acetate. The cell pellet was then resuspended in 100mM lithium acetate to give a final cell density of 2 X 10^9 cells/ml. The cell suspension was incubated at 30°C for 15 minutes and then dispensed into 50µl aliquots. Five microlitres of freshly boiled single-stranded DNA (10mg/ml) and up to 5µg of plasmid DNA were added to 50µl of lithium acetate treated cells. A 40% (w/v) polyethylene glycol/100mM lithium acetate solution (300µl) was added and vortexed thoroughly. The transformation mixture was incubated at 30°C for 30 minutes and then heatshocked at 42°C for 20 minutes. The cells were then centrifuged at 13000rpm for 10 seconds and the pellet was resuspended in 1ml of water. Aliquots (250µl) were plated onto the appropriate selective plates, air-dried and incubated at 30°C for 3-4 days.

2.2.24 Integration of linear DNA into the *S. cerevisiae* genome (Wach *et al.*, 1996)

A single colony of the appropriate strain of *S. cerevisiae* was used to inoculate 5ml of YPD and grown overnight at 30°C. The cell density of the pre-culture was estimated using a haemocytometer and diluted to 1 x 10^4 cells/ml in 50ml of YPD. The yeast cells were then grown at 30°C for

approximately 16 hours and then the cells were recounted (The cells were not used for the transformation if the cell density was higher than 1.7×10^7 cells/ml). The cells were collected by centrifugation at 3000r.p.m. (1500g) for 5 minutes using a GLC4 general laboratory centrifuge. The cells were washed once with 20ml of water and re-centifuged. The pellet was then resuspended in 1ml of water and transferred to a 1.5ml Eppendorf tube. The suspension was centrifuged at 5000r.p.m. for 5 minutes, using a MSE Micro-Ceutaur microfuge and the pellet was resuspended in 100mM lithium acetate to a final cell density of 2×10^9 cells/ml. These cells were then incubated for 20 minutes at 30°C to make them fully competent. Approximately 2.5 μg of the relevant purified DNA and 5μl of freshly boiled single-stranded salmon sperm DNA (10mg/ml) was added to 50μl of competent cells. The mixture was vortexed and incubated at 30°C for 20 minutes. A 40% (w/v) polyethylene glycol 3350/100mM lithium acetate solution (300μl) was added to each transformation mixture, vortexed and incubated again for 20 minutes at 30°C. The transformation mixture was then heat-shocked at 42°C for exactly 20 minutes and the cells were collected at 5000r.p.m. for 1 minute. The cell pellet was resuspended in 4 ml of YPD and incubated for 2-3 hours at 30°C. Aliquots (1.5ml) of each transformation culture were transferred to 1.5ml Eppendorf tubes, centrifuged at 5000r.p.m. for 1 minute and the cell pellet was resuspended in approximately 250μl of the remaining supernatant. The transformed cells were then plated onto YPD plates [supplemented with 200mg/l Geneticin (G418)] and incubated at 30°C for 2-3 days. The large colonies were then picked off the transformation plate, re-streaked onto a new YPD/G418 plate and incubated for 1-2 days.

2.3 PROTEIN METHODS

2.3.1 Sodium Dodecyl Sulphate PolyAcrylamide Gel Electrophoresis (SDS-PAGE)

Polyacrylamide gel electrophoresis was carried out under denaturing conditions (Laemmli, 1970) using the Bio-Rad Mini-Protean II electrophoresis system.

All samples were boiled for 5 minutes in 4x SDS loading buffer [62.5mM Tris-HCl (pH 6.8), 10% (v/v) glycerol, 2% (w/v) SDS, 0.1% (w/v) bromophenol blue, 870mM β-mercaptoethanol] and loaded onto a discontinuous polyacrylamide gel [6% (w/v) acrylamide stacking gel/10% (w/v) acrylamide running gel]. The 10% (w/v) acrylamide running gel was prepared by mixing: 3.33ml 30% (w/v) bis-acrylamide solution (Protogel™), 4.0ml water, 2.5 ml 1.5M Tris-HCl (pH8.8), 100µl 10% (w/v) SDS, 40µl 25% (w/v) ammonium peroxodisulphate and 4µl N, N, N', N'-tetramethylethylenediamine (TEMED). Approximately 3.5ml of the mixture was pipetted between the vertical glass plates of the Mini-Protean II system and 500µl of water was carefully added to ensure a smooth surface between the stacking and the running gel. Once set, the 6% (w/v) acrylamide stacking gel was prepared by mixing 1.7ml 30% (w/v) bis-acrylamide, 6.7ml water, 1.25ml 1M Tris-HCl (pH6.8), 100µl 10% (w/v) SDS, 40µl 25% (w/v) ammonium peroxodisulphate and 10µl TEMED. The water was removed from the top of the running gel, the stacking gel was added and a comb was inserted carefully to avoid any air bubbles. Once set, the SDS-PAGE gel was run at 200V for about 45 minutes in SDS running buffer [25mM Tris, 200mM glycine, 0.1% (w/v) SDS]. The gels were either fixed and stained simultaneously with 0.25% (w/v) Coomassie Brilliant Blue R250 [in 45% (v/v) methanol/10% (v/v) acetic acid] overnight at room temperature or transferred to nitrocellulose for Western analysis (Section 2.3.2). The gels were then destained with 30% (v/v) methanol/10% (v/v) acetic acid for approximately two hours at room temperature. After destaining, the gels were either stored in a sealed plastic bag in 5% (v/v) glycerol/5% (v/v) acetic acid or dried using a Bio-Rad 583 gel dryer.

2.3.2 Transfer of proteins from SDS-PAGE gels to nitrocellulose

Six pieces of Whatman 3MM filter paper and one piece of nitrocellulose were cut to the exact size of the SDS-PAGE gel. These were all soaked in transfer buffer [39mM glycine, 48mM Tris, 0.037%(w/v) SDS, 20%(v/v) methanol] for 20 minutes. Three pieces of pre-soaked filter paper were stacked onto the anode plate of the transfer apparatus (2117 Multiphore II electrophoresis unit, LKB Bromma) followed by the nitrocellulose, SDS-PAGE gel and finally another three pieces of the pre-soaked filter paper. The upper cathode plate was placed on top of the stack and a current of 0.8mA/cm^2 of gel was applied for 1.5 to 2 hours.

2.3.3 Detection of target protein via Western analysis

The nitrocellulose membrane was blocked using 2% (w/v) Marvel in phosphate-buffered saline [(PBS) 137mM NaCl, 2.7mM KCl, 9.2mM Na_2HPO_4, 1.8mM KH_2PO_4] for two hours at room temperature. The membrane was then washed twice, for 3 minutes each time, with PBS. The relevant primary antibody (Table 2.1) diluted suitably in PBS/2% (w/v) Marvel was incubated with the nitrocellulose membrane overnight (4°C) at 15 revolutions per minute on a 3D rocking platform (Stuart Scientific). The membrane was then washed twice, for 3 minutes each time, in PBS/0.05% (w/v) Tween 20 and then PBS. The relevant secondary antibody (Table 2.2), diluted appropriately in PBS/2% (w/v) Marvel, was then incubated with the membrane for 40 minutes at room temperature on the rocking platform. The PBS/0.05% (w/v) Tween 20 and PBS washes were repeated and the membrane was developed using the chromogenic substrate 4-chloro-1-naphthol [0.6% (w/v) 4-chloro-1-naphthol, 20% (w/v) methanol, 8mM Tris-HCl (pH7.5), 120mM NaCl, 0.015% (v/v) hydrogen peroxide]. After development, the membrane was fixed using water, dried and then stored in a sealed plastic bag.

Table 2.1 - Primary antibodies

Supplier	Product code	Host animal	Antibody specificity	Working dilution	Amount used (μg/ml)
Sigma	I9010	Goat	human IgG Fab specific	1/1000	4
Sigma	M0614	Mouse	human lambda light chain	1/100	2.35
Dako	A191	Rabbit	human kappa light chain	1/800	0.75
Dako	A0423	Rabbit	human IgG, γ-chain specific	1/500	2.5
Gift from D.Jeenes		Rabbit	hen egg white lysozyme	1/2000	
Jackson	315-005-006	Rabbit	mouse IgG, $F(ab')_2$ specific	1/1000	2.3

Table 2.2 - Secondary antibodies

Supplier	Product Code	Host animal	Antibody specificity	Working Dilution
Sigma	A0545	Goat	rabbit IgG, peroxidase conjugate	1/1000
Sigma	A4416	Goat	mouse IgG. peroxidase conjugate	1/1000
Sigma	A5420	Goat	goat IgG, peroxidase conjugate	1/1000
Jackson	111-095-144	Goat	rabbit IgG, FITC conjugate	1/100

2.3.4 Enzyme- Linked Immunosorbent Assay (ELISA)

The wells of an ELISA plate were coated with 100µl of the relevant antigen, diluted appropriately in PBS and were left overnight at room temperature. The wells were rinsed three times with PBS and then blocked for 2 hours at 37°C, using 2% (w/v) Marvel/PBS. The wells were then washed three times with PBS and 25µl of 10% (w/v) Marvel/PBS was added to each well. The *E. coli* supernatant (100µl) from the appropriate culture was added to the designated well and left for 1.5 hours at room temperature. The wells were then rinsed three times with PBS/0.05% (w/v) Tween 20 and then three times with PBS. A hundred microlitres of the relevant primary antibody (Table 2.1), diluted in PBS/2% (w/v) Marvel, was added to each of the wells and left at room temperature for 1 hour. The PBS/0.05% (w/v) Tween 20 and PBS washes were repeated. A hundred microlitres of the complementary secondary antibody (Table 2.2) diluted in PBS/2% (w/v) Marvel was added to each well and incubated at room temperature for 30 minutes. The PBS/0.05% (w/v) Tween 20 and PBS washes were repeated and the chromogenic substrate 3, 3', 5, 5'-tetramethyl benzidine [TMB (50mM citrate buffer (pH6.0), 4mM TMB, 2 x 10^{-4} % (v/v) hydrogen peroxide)] was used to develop the wells. The reaction was stopped using 50µl of 1M sulphuric acid and the adsorbance (450nm) was read using a Multiskan® MCC Labsystems ELISA plate reader.

2.3.5 Enzyme ImmunoAssay (EIA)

Approximately 2 x 10^7 yeast cells were added to a 1.5ml Eppendorf tube and washed three times with PBS. Two hundred microlitres of the primary antibody (Table 2.1) diluted in PBS/2% (w/v) Marvel was added, vortexed and left for 1 hour at room temperature on a vertically rotating platform. The cells were collected via centrifugation at 13000r.p.m. for 30 seconds on a MSE Micro-Ceutaur microfuge. The cell pellet was washed three times with PBS/0.05% (w/v) Tween 20 and then three times with PBS. Two hundred microlitres of the complementary secondary antibody (Table 2.2) diluted in 2% (w/v) Marvel/PBS was added, vortexed and left for 30 minutes at room temperature on the rotating platform. The washes were then repeated and a

100µl of the chromogenic substrate 3, 3', 5, 5'-tetramethyl benzidine (TMB) [50mM citrate buffer (pH6.0), 4mM TMB, 2 x 10^{-4} % (v/v) hydrogen peroxide] was added to the cell pellet. The reaction was stopped using 50µl of 1M sulphuric acid and re-centrifuged. The reaction supernatant (100µl) was transferred to an ELISA plate and the absorbance (450nm) was read using a Multiskan® MCC Labsystems ELISA plate reader.

PART I

MIMICKING ANTIBODY-DISPLAYING B-CELLS USING *SACCHAROMYCES CEREVISIAE*

Strategy

During the germinal centre reaction (Section 1.3.3), B-cells display mutated antibodies on their cell surfaces. This part of the project is concerned with copying the display system of B-cells by over-expressing antibody fragments on the surface of *S. cerevisiae* and determining if these antibody fragments are accessible to the relevant antigen. The expression constructs used in this project were designed with the intention of developing a fluorescence-activated cell sorting (FACS) selection system that could distinguish between yeast cells displaying low or high affinity antibodies for a particular antigen on their surfaces. The proposed FACS selection system was devised to imitate the antibody selection process of the germinal centre reaction.

Numerous heterologous proteins have been covalently attached to the β-glucan component of the cell wall of *S. cerevisiae* using the C-terminal anchorage domain of the α-agglutinin gene (Section 1.7). The α-agglutinin anchorage domain contains a glycosyl phosphatidyinositol (GPI) attachment sequence that is essential for linkage to the β-1,6 glucan component of the cell wall (Section 1.6.2). This α-agglutinin anchorage domain will be used to display antibody fragments (scFv and Fab molecules) on the cell surface of *S. cerevisiae.*

When designing the strategy for high level expression of antibody fragments on the surface of *S. cerevisiae,* the following factors were taken into consideration (Romanos *et al.*, 1992):

1. The copy number of the antibody gene.
2. The nature of the promoter.
3. The type of signal sequence.
4. The stability of the mRNA.

5. The stability of the protein after synthesis.

Firstly, the expression vectors used in this part of the project contain a fragment from the multi-copy plasmid present in most laboratory strains known as the 2μ circle (Beggs, 1978). This fragment gives the vectors the ability to replicate autonomously and be stably maintained at 25-100 copies per cell. Additionally, the vectors carry an *E. coli* selectable marker and origin of replication. These will facilitate the cloning of the antibody fragments into the expression vectors using standard DNA manipulation techniques.

High level expression of a foreign gene can place a significant burden on the host cell, reducing its growth rate and affecting the efficiency of gene expression (Romanos *et al.*, 1992). In constitutive systems, where growth and expression are linked, strong selective pressure exists for cells with reduced levels of foreign gene expression. To overcome this problem, this strategy employs a strong yeast inducible promoter (*GAL1*; Section 3.1.1) which can largely separate the growth and expression phases.

To target foreign proteins through the yeast secretory pathway, a signal sequence must be used which normally consists of a charged N-terminus, a central hydrophobic core and a consensus sequence for cleavage in the endoplasmic reticulum. In this case, the native α-agglutinin signal sequence was fused to the 5' terminus of the antibody fragments to direct their expression to the secretory pathway (Section 1.6.2). Additionally, a yeast transcriptional termination sequence was placed at the 3' terminus of each construct to aid efficient 3' formation of the messenger RNA (Romanos *et al.*, 1992).

Finally, all expression vectors were transformed into a protease-deficient *S. cerevisiae* strain to overcome the severe protease problem in yeast (Jones, 1991). The lumenal vacuolar aspartyl (prA) and serine (prB) proteases are thought to be the major source of this problem, therefore the *S. cerevisiae* strain used in this part of the project was chosen to be deficient in these. This should decrease the amount of protein degradation that can occur after cell breakage.

The single chain Fv and Fab antibody fragments expression constructs consist of an α-agglutinin signal sequence, the C-terminal anchorage domain of the α-agglutinin gene (except for the case of the light chain Fab construct) and a transciptional termination signal. Each construct was placed under the control of the inducible *GAL1* promoter (Figure 3.3).

The light and heavy chains of the Fab antibody fragment were cloned into separate expression vectors which comprised of the same promoter and signal sequence but different yeast auxotrophic markers for individual selection. Hence, it is intended that the light and heavy chains will be expressed simultaneously through the secretory pathway and the Fab fragment will form in the endoplasmic reticulum where disulphide bond formation takes place. The α-agglutinin anchorage domain fused to the heavy chain will hopefully attach the assembled Fab fragment to the cell wall. Functional Fab fragments have been successfully assembled from separate expression vectors under the control of the same promoter and secreted into the culture supernatant (Horwitz *et al.*, 1988).

A Factor Xa cleavage site (Ile-Glu-Gly-Arg; Nagai and Thogersen, 1984) was placed in-between the antibody fragment and the α-agglutinin anchorage domain in both the scFv and heavy chain surface display constructs. Therefore, antibody fragments could be detached from the cell wall with treatment with Factor Xa protease, which cuts at a single arginine residue at the cleavage site. This may aid in the characterisation of the antibody fragment affinities using techniques such as fluorescence quench titration or surface plasmon resonance. This part of the strategy is intended to mimic the soluble antibody producing plasma cells of the human immune system.

Once surface expression of the antibody fragments is established, it is proposed that yeast cells displaying antibodies with high affinity for a particular antigen could be selected from a mixed population of cells using fluorescence-activated cell sorting (FACS). The antibody fragments used in this part of the project bind to the fluorescent antigen, fluorescein 5-isothiocyanate, therefore the feasibility of the FACS selection system could be shown directly.

CHAPTER 3

CONSTRUCTION OF A YEAST SURFACE DISPLAY SYSTEM FOR ANTIBODY FRAGMENTS

3.1 INTRODUCTION

The first aim of this project is to show that antibody fragments (scFv and Fab molecules) can be functionally expressed on the surface of *Saccharomyces cerevisiae*. This strategy is intended to mimic the antibody surface display system of B-cells in the germinal centre reaction (Section 1.3.3).

The C-terminal anchorage domain of the α-agglutinin gene has been used to display several heterologous proteins on the surface of *S. cerevisiae* (Section 1.7) and is the chosen expression system for this part of the project (Part 1). The α-agglutinin anchorage domain will be fused to the C-terminus of the antibody fragments. Therefore, when the antibodies are expressed they will hopefully become covalently attached to the β-glucan component of the cell wall and be displayed on the exterior surface of *S. cerevisiae*.

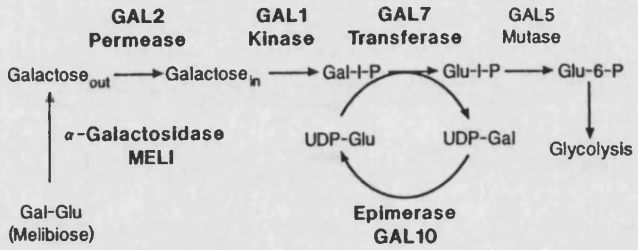
The human antibody fragments used in this section of the project (Chapter 3 and 4) bind to the fluorescent antigen, fluorescein 5-isothiocyanate (FITC; Griffiths *et al.*, 1994). It is proposed that by using anti-FITC antibody fragments, epi-fluorescence microscopy can be used directly to determine if functional antibodies have been expressed on the surface of *S. cerevisiae*. Additionally, cells binding the fluorescent antigen could be individually selected using fluorescence-activated cell sorting (FACS). This selection method could be used to isolate high affinity antibodies from a expression library of mutated antibodies, analogous to the germinal centre reaction.

The background information concerning the inducible *GAL1* promoter and the anti-FITC Fab fragments is reviewed below.

3.1.1 The inducible promoter, *GAL1*

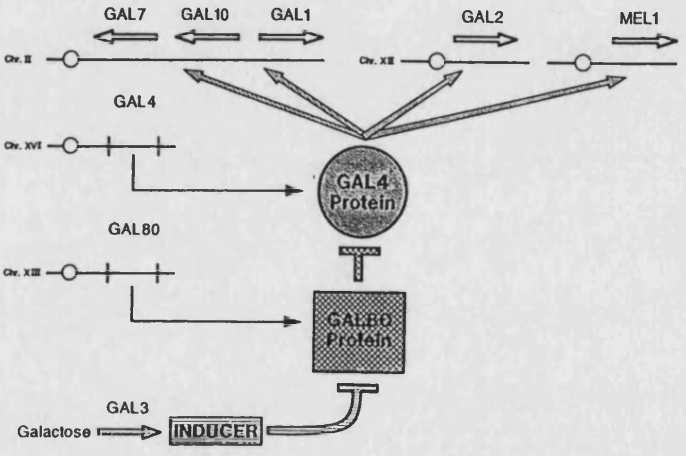
The mechanism of regulation of the enzymes used in galactose utilization (*GAL* genes) is probably one of the most well characterised in yeast (Johnston, 1987). Galactose is converted to glucose-6-phosphate by the enzymes of the Leloir pathway (Figure 3.1). The genes: *GAL2* (permease), *GAL1* (kinase), *GAL7* (transferase), *GAL10* (epimerase) and *GAL5* (mutase) encode the enzymes of this pathway.

Figure 3.1 – The pathway of galactose utilization. Reproduced from Johnston, 1987.



Expression of these enzymes (except for *GAL5*) is tightly regulated and increases approximately one thousand fold after growth on galactose. Transcription of these genes is activated by a protein, GAL4, which binds to sequences located upstream from each of these genes [upstream activation sequences (UAS)]. Growth on galactose causes an inducer to bind to the GAL80 protein. This prevents the inhibition of the GAL4 protein function (Figure 3.2) and transcription is activated.

Figure 3.2 – The components of the *GAL* gene regulatory circuit. Reproduced from Johnston, 1987.



Using glucose as a carbon source prevents *GAL* gene expression by an event termed catabolite repression. Little is known about the catabolite repression mechanism but even in the presence of galactose, expression of *GAL* genes is completely repressed at the level of transcription. Cells grown on a non-catabolite-repressing carbon source such as raffinose require only ten to twenty minutes after galactose addition to induce transcription of the *GAL* genes (Adams, 1972). Cells grown in glucose display a delayed induction time of approximately 3 to 5 hours, even after extensive washing to remove any excess glucose.

In this project, the *GAL* promoter consists of the divergent upstream activation sequence from the promoter region between the *GAL1* and *GAL10* genes. Galactose induction, after growth in a non-catabolite repressing carbon source, should display a 1000-fold increase in expression levels.

3.1.2 The human anti-FITC antibody fragments

Numerous different anti-FITC Fab fragments with different human light and heavy gene segments (and randomised CDR3s) were isolated from a very large synthetic phage display library created by 'combinatorial infection and *in vivo* recombination' (Griffiths *et al.*, 1994; Section 1.4.4). This part of the project utilizes two of these isolated human anti-FITC Fab fragments (A2 and B7), both of which have affinities comparable with antibodies from a secondary immune response (Table 3.1).

Table 3.1 – The two human anti-FITC Fab antibody fragments used in this part of the project (Griffiths *et al.*, 1994). The residues in the light chain CDR3 regions encoded by randomized codons are in bold.

Clone	Heavy chain	CDR3	Light chain	CDR3	Kd (nM)
FITC-A2	DP-47	YRFSAPPRD	DPL-3	AAWDDSL**P**S**G**V	217(± 16)
FITC-B7	DP-67	AQRKYFDY	DPK-12	MQSIQL**R**T	151(± 3)

3.1.3 Experimental design for the construction of the surface display expression vectors

The first stage in the construction of the anti-FITC Fab expression vectors (pYES2-*SAG1*-Heavy A2 and pYX-Light A2; Figure 3.6 and 3.8) was to individually amplify the heavy and light Fab gene segments from the plasmid, pUC119His6mycXba-FITC A2. The 5' oligonucleotides used in both the amplifications were designed to encode for the α-agglutinin signal sequence. The α-agglutinin signal sequence was incorporated into all the expression constructs to target the antibody fragments through the yeast secretory pathway (Section 1.6.2). A restriction endonuclease recognition site (*Msc* I) was introduced at the 3' end of the α-agglutinin signal sequence to aid the cloning of other antibody gene fragments.

The second stage of construction consisted of amplifying the C-terminal anchorage domain of the α-agglutinin gene and fusing this fragment to the 3' end of the Fab heavy chain gene segment using a 'splicing by overlap extension' reaction (Horton *et al.*, 1989). The complementary oligonucleotides used in this reaction also encoded for a Factor Xa cleavage site and a *Not* I recognition site. The Factor Xa cleavage site was present to detach the scFv or Fab molecules from the yeast cell wall (if desired) and the *Not* I recognition site was introduced to facilitate the cloning of other antibody gene fragments.

The yeast expression vectors chosen for this project, pYES2 and pYX243, comprised of a multiple cloning site flanked by a strong *GAL1* inducible promoter and a transcription termination sequence (Section 3.2.3b). Careful design of the oligonucleotides facilitated the cloning of the light chain gene and heavy chain-α-agglutinin gene fusion into the multiple cloning sites of the appropriate expression vectors using restriction endonucleases. The vectors used for the surface display of the anti-FITC Fab A2 molecules contained different yeast auxotrophic selectable markers. Hence, the individual heavy and light chain expression constructs were co-transformed into an appropriate protease-deficient yeast strain and simultaneously expressed due to the fact they were under the control of the same *GAL1* inducible promoter. It was reasoned that the separate heavy and light chains

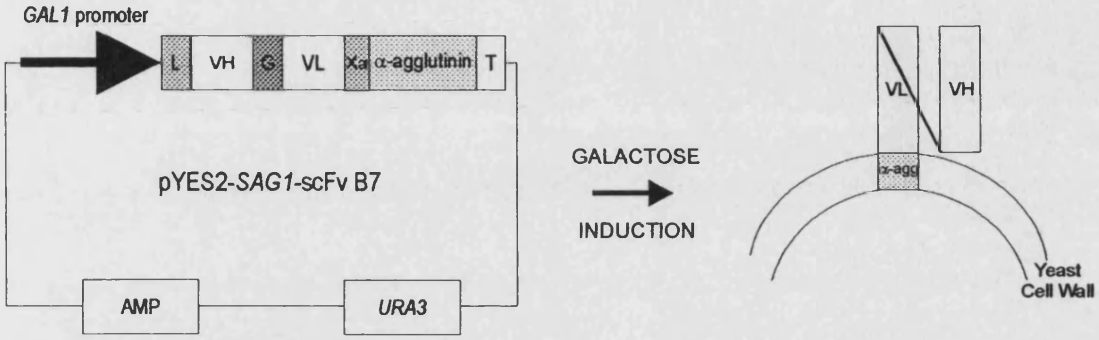
would form a functional Fab fragment in the endoplasmic reticulum and the α-agglutinin anchorage domain, which is fused to the C-terminus of the heavy chain, would covalently attach the Fab to the cell wall.

The construction of the anti-FITC scFv surface display vector (pYES2-*SAG1*-scFv B7) involved separately amplifying the heavy and light variable gene segments from the plasmid, pUC119His6mycXba-FITC B7 (Figure 3.10). The variable heavy and light gene segments were then fused together by a 'splicing by overlap extension' reaction which also incorporated a $(Gly_4Ser)_3$ linker (Huston *et al.*, 1988). Due to the careful design of the heavy chain expression vector, the construction of the scFv surface display vector was straightforward. Firstly, the heavy chain was removed from the pYES2-*SAG1*-Heavy A2 by using the two restriction endonuclease recognition sites (*Msc* I and *Not* I) placed at the 3' ends of the α-agglutinin signal sequence and heavy chain gene segment. The 5' and 3' ends of the scFv gene were designed to be compatible with these recognition sites, hence the scFv was digested and ligated into this expression vector. The created expression vector, pYES2-SAG1-scFv B7, consisted of the scFv gene flanked by an α-agglutinin leader sequence and the α-agglutinin anchorage domain, all under the control of the *GAL1* promoter.

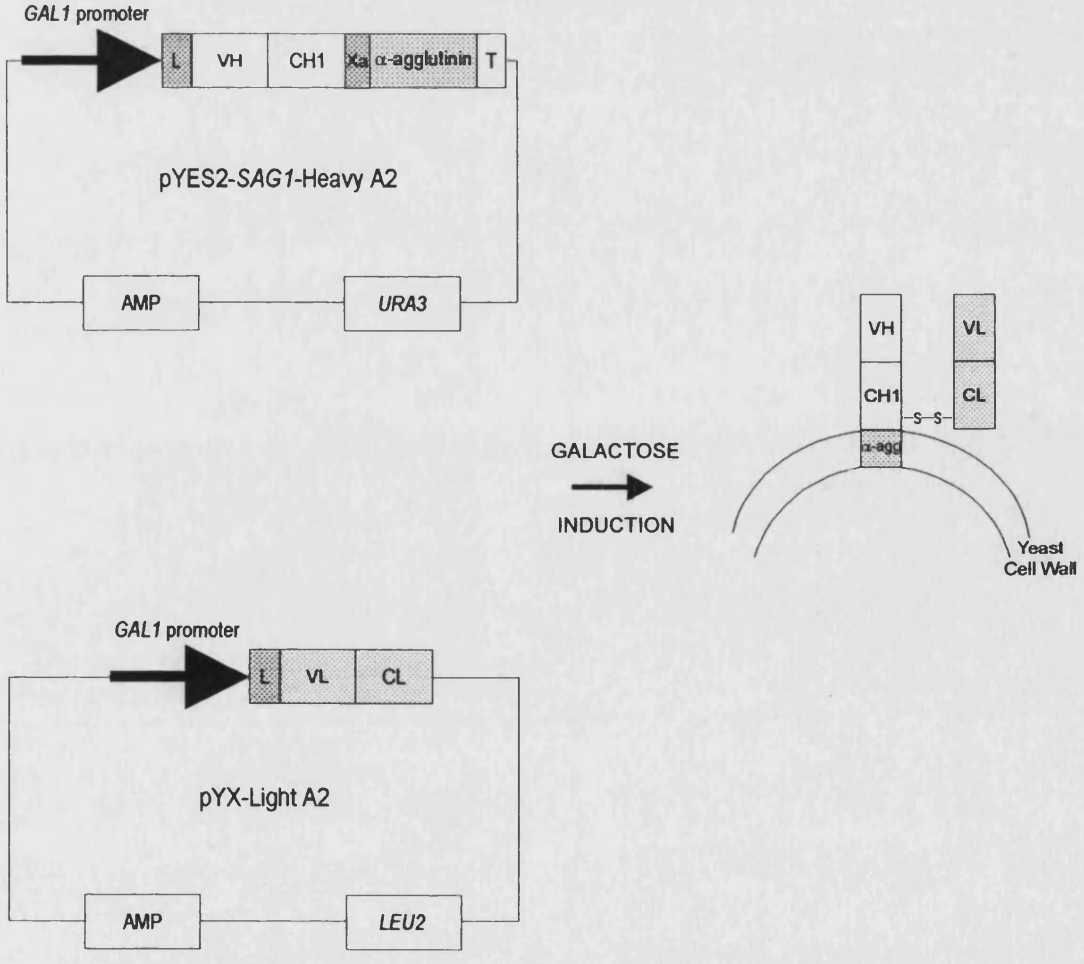
This Chapter details the successful construction of anti-FITC Fab and scFv surface display expression vectors, pYES2-*SAG1*-Heavy A2, pYX-Light A2 and pYES2-*SAG1*-scFv B7 (Figure 3.3).

Figure 3.3 - Diagrammatic representation of the constructed surface expression vectors. The letters L, G, T and Xa represents the α-agglutinin leader sequence, $(Gly_4Ser)_3$ linker, termination sequence and the Factor Xa recognition site, respectively.

a) The anti-FITC single chain Fv expression vector, pYES2-*SAG1*-scFv B7.



b) The surface display vectors, pYES2-*SAG1*-Heavy A2 and pYX-Light A2, designed for expression of the anti-FITC Fab A2 antibody fragment.



3.2 MATERIALS

3.2.1 *Saccharomyces cerevisiae* strain

The *S. cerevisiae* strain, YMTAB contained gene disruptions in the vacuolar proteases, pra1 and prb1 (Jones, 1991). The strain was galactose inducible and contained the relevant auxotrophic mutations required for plasmid selection.

Table 3.2 – *Saccharomyces cerevisiae* strain used.

Strain	Genotype	Source	Reference
YMTAB	*MATα leu2-3,112 ura3Δ5 pra1::HIS3 prb1ΔAV GAL⁺*	D. H. Wolf, University of Stuttgart	Sander *et al.*, 1994

3.2.2 *E. coli* strains and plasmids

The plasmids used in this chapter were either supplied in an aqueous DNA solution or as an *E. coli* slant culture (Table 3.3). When plasmids were provided in an aqueous solution, they were transformed into TOP10F' *E. coli* cells (Section 2.2.15). For long term storage, glycerol stocks were made for all plasmids (Section 2.2.3).

3.2.3 Plasmid descriptions

A) The pUC119His6mycXba–FITC A2/FITC-B7 plasmids (Griffiths et al., 1994)

The pUC119His6mycXba expression vector contains the anti-FITC heavy and light Fab chains cloned into the plasmid as *Xba I/Not* I fragments. The tags, c-myc and His^6, are present after the *Not* I site for detection and purification of the anti-FITC Fab fragments, respectively. The Fab fragments are under the control of the inducible promoter, *lac,* and therefore expression can be induced by addition of isopropyl-β-D-thiogalactoside (IPTG).

B) The yeast expression vectors, pYES2 and pYX243

The vectors, pYES2 and pYX243, contain selectable markers and origins of replication for *S. cerevisiae* and *E. coli*, therefore can be maintained in both (Figures 3.4 and 3.5). They are designed for high level expression of foreign proteins in *S. cerevisiae* and carry a strong inducible *GAL1* promoter and transcription terminator sequence.

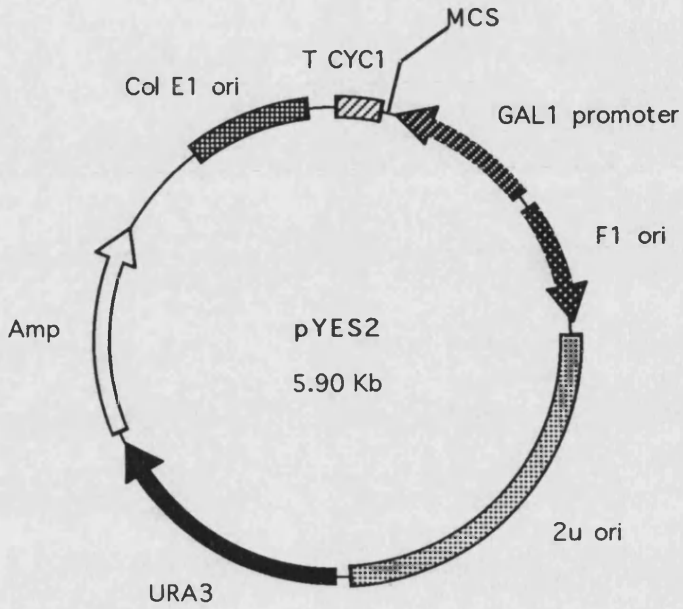
C) The pH27 vector (Lipke et al., 1989)

The *E. coli* cloning vector, pH27, carries a 6.1k.b.p. *SAG1*-complementing *Hind* III fragment which contains the complete α-agglutinin gene. A kanamycin marker and origin of replication are present for the selection and replication in *E. coli,* respectively.

Table 3.3 – Plasmids used. NA, not applicable.

Plasmid	Strain	Genotype	Source	Reference
pUC119His6mycXba -FITCA2	TG1	*supE hsdΔ5 thiΔ(lac⁻ proAB) F' (traD36 proAB⁺ lacI^q lacZΔM15)*	G. Winter, MRC, Cambridge	Griffiths *et al.*, 1994
pUC119His6mycXba -FITC B7	TG1	" "	" "	" "
pYES2	NA	NA	Invitrogen	NA
pYX243	NA	NA	R + D systems	NA
pH27	NA	NA	H. de Nobel, University of Vermont	Lipke *et al.*, 1989

Figure 3.4 - Plasmid map of the yeast expression vector, pYES2.

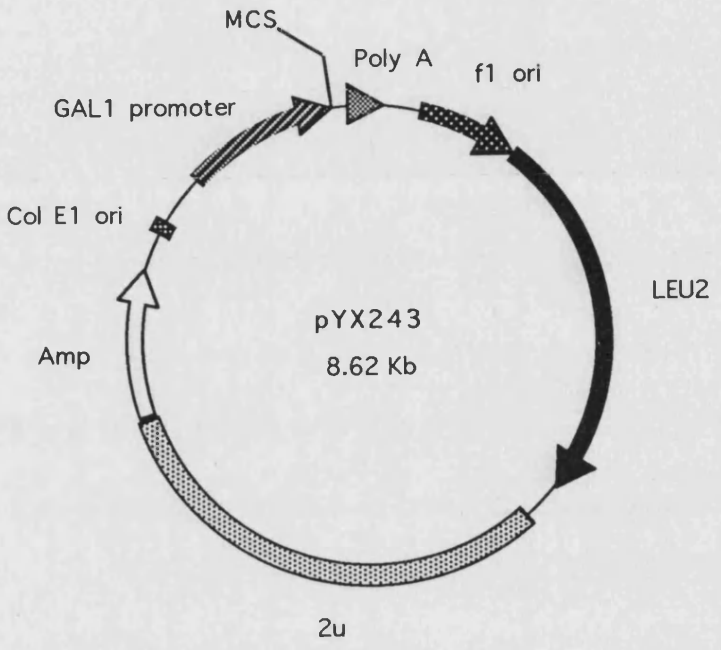


Multiple cloning site (MCS) for the expression vector, pYES2.

Xba I *Sph* I *Xho* I *Not* I *BstX* I *EcoR* I *BstX* I
TCTAGATGCATGCTCGAGCGGCCGCCAGTGTGATGGATATCTGCAGAATTCCAGCACA

*Bam*H I *Sac* I *Kpn* I *Hind* III
CTGGCGGCCGTTACTAGTGGATCCGAGCTCGGTACC**AAGCTT**

Figure 3.5 - Plasmid map of the yeast expression vector, pYX243.



Multiple cloning site for expression vector, pYX243

BamH I *Apa I* *Sal I*
EcoR I *Nco I* *Avr II* *Hind III* *Mlu I* *Sma I*
GAATTCACCATGGATCCTAGGGCCCACAAGCTTACGCGTCGACCCGGGTATCCGTATG

Xho I
EcoR V *Sac I* *Nhe I*
ATGTGCCTGACTACGCATGATATCTCGAGCTCGAGCTCA**GCTAGC**TAACTGA

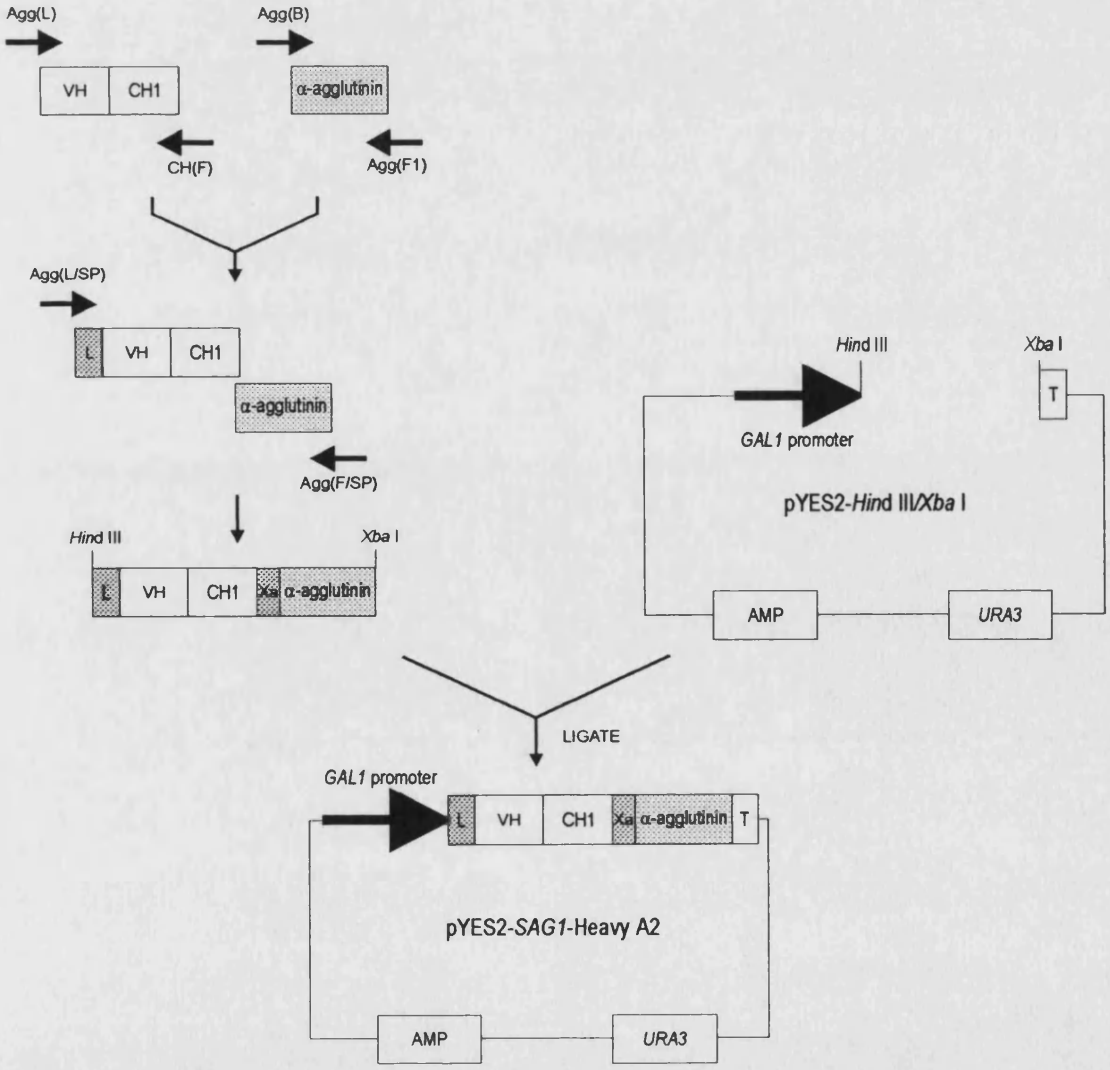
3.3 METHODS

3.3.1 Construction of the Fab heavy chain expression vector

A) *Primer design*

The fundamental design parameters for primers were followed as described in Section 2.2.4. Primer pairs, Agg(F1)/Agg(B) and Agg(L)/CH(F) were designed to amplify the C-terminal half of the α-agglutinin gene and the anti-FITC Fab A2 heavy chain (germline sequence DP-47 and the IgG constant domain 1), respectively (Figure 3.6).

Figure 3.6 – Flow diagram summarising the construction of the Fab heavy chain expression vector, pYES2-*SAG1*-Heavy A2. The letters L and Xa depict the α-agglutinin leader sequence and Factor Xa recognition site, respectively. The symbols VH and CH1 represents the variable heavy chain gene segment (DP-47) and the first constant domain of an IgG1 molecule, respectively.



Oligonucleotides Agg(B) and CH(F) were synthesised with complementary 5' ends. Therefore, in a second PCR step, the heavy chain-α-agglutinin fusion can be assembled by a 'splicing by overlap extension' reaction (Horton *et al.*, 1989). These complementary primers also encoded for a Factor Xa cleavage site to aid the removal of antibody fragments from the yeast cell wall, if desired. To facilitate cloning *Hind* III and *Xba* I recognition sites were incorporated into the 5' ends of the oligonucleotides, Agg(L) and Agg(F1), respectively. Additionally, the primer Agg(L) was designed to encode for the α-agglutinin signal sequence .

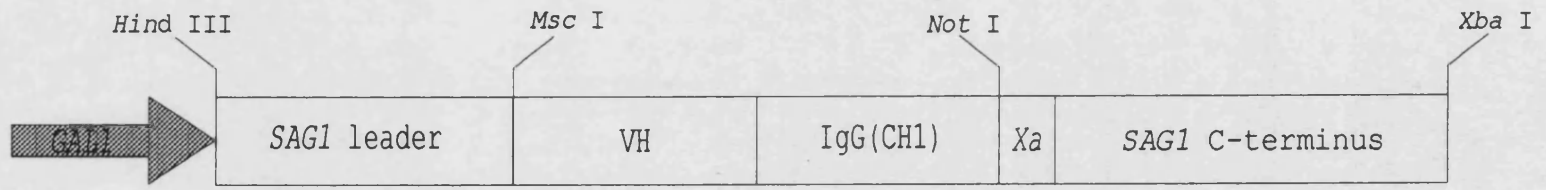
In anticipation, the heavy chain surface display construct would be used to clone additional antibody fragments, primers Agg(L) and CH(F) were designed to incorporate a *Msc* I and a *Not* I site at the 3' end of the α-agglutinin leader sequence and 5' end of the α-agglutinin anchorage domain, respectively. This should lead to the easy removal and addition of antibody gene fragments in-between the α-agglutinin leader and anchorage domain. Figure 3.7 summarizes the heavy chain A2/α-agglutinin expression construct.

B) *PCR amplification of the C-terminal α-agglutinin*

The C-terminal anchorage domain of the α-agglutinin gene (nucleotides 982-1952) was amplified by a $Vent_R$ PCR (Section 2.2.7a) using the vector, pH27 as a DNA template and the primer pair, Agg(F1) and Agg(B) [Table 3.4]. The reaction was performed for 25 cycles using the following parameters: [(94°C, 1min; 50°C, 2min; 72°C, 1min) x 25; 72°C, 10min]. The PCR product (986 b.p.) was purified (Section 2.2.9b) and quantified by agarose gel electrophoresis (Section 2.2.8).

Note: M. Redfern optimised this PCR amplification.

Figure 3.7 - Schematic representation of the heavy chain A2/α-agglutinin gene expression construct including sequence details of the 5' and 3' ends. The DNA sequence encoding the heavy A2/α-agglutinin junction is also described. Restriction sites are represented in bold and the extra nucleotide is underlined.



5' end

α-agglutinin leader sequence

1 10 20

M F T F L K I I L W L F S L A L A S A E V Q L L E

AAGCTTATGTTCACTTTTCTCAAAATTATTCTGTGGCTTTTTTCCTTGGCAT**TGGCCA**GTGCTGAGGTGCAGCTGTTGGAG

Hind III *Msc I*

VH (DP-47)

Heavy chain A2/α-agglutinin junction *3' end*

CH1/Hinge α-agglutinin gene α-agglutinin gene

234 240 250 565 570

V E P K S C A A A I E G R G T A S A K L S Y L L F O O

GTTGAGCCCAAATCTTGT**GCGGCCGC**CATCGAGGGAAGGGGTACAGCTAGCGCCAAA TTGTCGTACCTGCTATTCTAATAA**TCTAGA**

Not I *Xba I*

Factor Xa cleavage site

C) *PCR amplification of the Fab heavy chain A2*

The human heavy chain A2 consisting of the variable gene segment DP-47 and first IgG heavy constant domain was amplified from vector, pUC119His6mycXba-FITC A2 using an Expand™ PCR (Section 2.2.7c) and primers Agg(L)/CH(F) [Table 3.4].The reaction conditions were: [(94°C, 1min; 60°C, 1min; 72°C, 1min) x 25; 72°C, 10min] and the PCR product (749 b.p.) was purified using a DNA extraction kit (Section 2.2.9b).

Table 3.4 – Sequences of the oligonucleotides used in the amplification of the 3' half of the α-agglutinin gene and the heavy chain A2 gene segment. Enzyme recognition sites are represented in bold and the Factor Xa site is underlined.

Primer Name	Sequence (5'-3')	Description (5'-3')
Agg(B)	ATCGAGGGAAGG**GGTACC**GCTAGCGCC AAA	12 n.t. Factor Xa site; binds 982-999 b.p. downstream from the *SAG1* start codon (Watson strand). The first six n.t. of homology to *SAG1* also incorporates a *Kpn* I site.
Agg(F1)	GC**TCTAGA**TTATTAGAATAGCAGGTACG ACAA	2 n.t. overhang; a *Xba* I site; one extra stop codon; 21n.t. of homology to the 3' end of *SAG1* (Crick strand).
Agg(L)	CCCCC**AAGCTT**ATGTTCACTTTTCTCAAA ATTATTCTGTGGCTTTTTTCCTTGGCATT **GGCCA**GTGCTGAGGTGCAGCTGTTGGA G	5 n.t overhang; a *Hind* III site; 57 n.t. α-agglutinin signal sequence including an *Msc* I site; 18 n.t. of homology to the 5 ' end of the germline antibody sequence, DP-47 (coding strand).
CH(F)	GCTAGC**GGTACC**CCTTCCCTCGATG**GC GGCCGC**ACAAGATTT	24 n.t. which are complementary to Agg(B) and encompasses a 12 n.t. Factor Xa site; one extra guanine n.t.; 18 n.t. of homology to the 3' end of the first IgG constant domain including a *Not* I site (non-coding strand).

D) *Assembly of the heavy A2/α-agglutinin fusion*

Approximately equal amounts (~50ng) of the amplified heavy chain A2 and the α-agglutinin anchorage domain were added together in a $Vent_R$ PCR (Section 2.2.7a) with the primer pair Agg(L/SP)/Agg(F/SP) [Table 3.5]. The reaction conditions were: [(94°C, 1min; 55°C, 2min; 72°C, 1min) x 25; 72°C, 10min]. The assembled product (1.7 k.b.p.) was separated from the other non-specific bands by low-melting point agarose gel electrophoresis (Section 2.2.8) and recovered using a DNA extraction kit (Section 2.2.9b).

E) *Construction of the vector, pYES2-SAG1-Heavy A2*

The vector, pYES2 and the heavy A2-α-agglutinin gene fusion were digested with *Hin*d III (10units) and *Xba* I (20units) overnight, at 37°C (Section 2.2.12). The digested pYES2-*Hin*d III/*Xba* I vector was separated from its multiple cloning site using low-melting point agarose gel electrophoresis (Section 2.2.8) and purified (Section 2.2.9b). Additionally, the restriction digestion buffers from the digested heavy A2/α-agglutinin fusion were removed (Section 2.2.9b). The heavy A2/α-agglutinin–*Hin*d III/*Xba* I insert was ligated into plasmid, pYES2-*Hin*d III/*Xba* I (Section 2.2.13) and transformed into *E. coli* (Section 2.2.15). Ampicillin-resistant colonies were PCR screened (Section 2.2.16) using conditions: [(94°C, 1min; 50°C, 1min; 72°C, 2min) x 25; 72°C, 10min] and the primer pair pYES(FOR2)/pYX(FOR) [Table 3.5]. The PCR products were analysed by agarose gel electrophoresis (Section 2.2.8) and positive transformants were grown overnight in 3ml of LB (Section 2.2.1). Glycerol stocks (Section 2.2.3) and plasmid DNA (Section 2.2.17) were prepared for some of the positive colonies. The construction of vector, pYES2-*SAG1*-Heavy A2 was verified by DNA sequencing (Section 2.2.18) using primers pYX(FOR), Agg(B), AggSeq1 and pYES(FOR2).

Table 3.5 - Sequences of the oligonucleotides used in the construction and verification of the heavy chain expression vector, pYES2-*SAG1*-Heavy A2. Restriction enzyme recognition sites are shown in bold.

Primer Name	Sequence (5'-3')	Description (5'-3')
Agg(L/SP)	CCCCC**AAGCTT**ATGTTCA	18 n.t. of homology to the 5' end of oligonucleotide Agg(L) including a *Hind* III site (coding strand).
Agg(F/SP)	GC**TCTAGA**TTATTAGAAT	18 n.t. of homology to the 5' end of primer Agg(F1) including a *Xba* I site (non-coding strand).
PYES(FOR2)	TTCGGTTAGAGCGGATGT	Binds 51-69 n.t. downstream of the *Xba* I site in pYES2 (coding strand).
PYX(FOR)	TACTTTAACGTCAAGGAG	Binds 57-75 n.t. upstream of the *Hind* III site on pYES2 (non-coding strand).
AggSeq1	GCCTGTTTCTACTGTGGAAAT	Binds 1068-1089 b.p. from the *SAG1* start codon (crick strand).

3.3.2 Construction of the Fab light expression vector

A) *Primer design*

The oligonucleotide Agg(L/L) was designed to anneal to the 5' end of the human variable lambda light chain (DPL-3) and also encoded the α-agglutinin leader sequence (Figure 3.8). The oligonucleotide CL(FORX) was designed to bind to the 3' end of the human lambda constant region. To clone the amplified light chain A2 into the digested expression vector, pYX243-*Eco*R I/*Nhe* I, the primers Agg(L/L) and CL(FORX) were designed with restriction sites, *Eco*R I and *Xba* I, respectively. The *Nhe* I recognition site could not be used because there was an unwanted *Nhe* I site in the lambda constant region. Therefore, the restriction site *Xba* I was recruited which produces compatible cohesive ends with *Nhe* I, when digested.

Ligation of the *Nhe* I and *Xba* I cut sites removes both the recognition sites. Therefore to facilitate the removal and addition of the light chain gene segment, an *Asc* I recognition site was placed upstream from the *Xba* I site. Figure 3.9 details the design of the light A2 expression construct.

Figure 3.8 – Flow diagram describing the construction of the Fab light expression vector, pYX-Light A2. The letters L, VL and CL depict the α-agglutinin leader sequence, variable light chain gene segment (DPL-3) and the human lambda constant region, respectively.

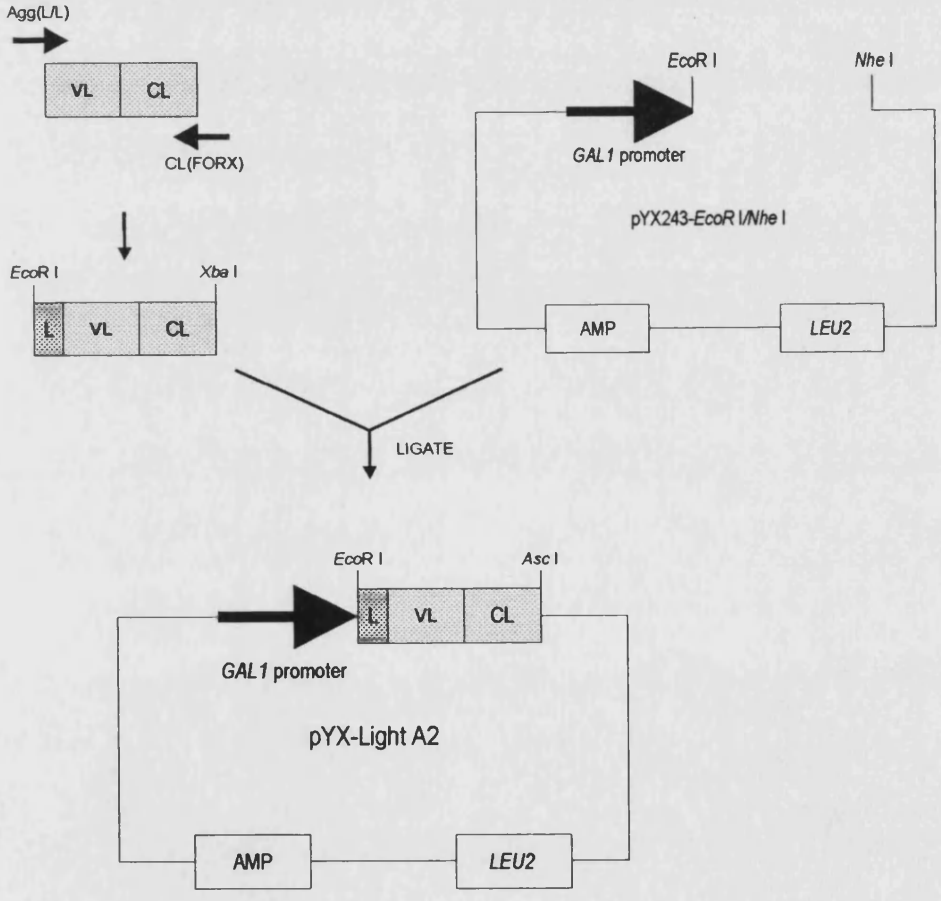
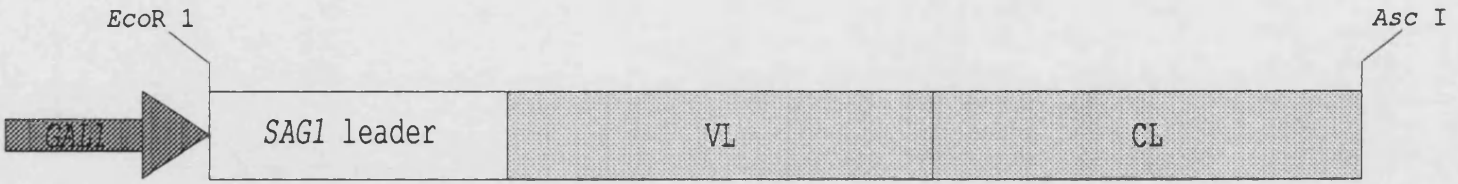


Figure 3.9 - Schematic representation of the light chain A2 construct including sequence details of the 5' and 3' ends. The restriction sites are represented in bold.



5' end

α-agglutinin leader sequence

1 10 20

M F T F L K I I L W L F S L A L A S A Q S V L T Q

GAATTCATGTTCACTTTTCTCAAAATTATTCTGTGGCTTTTTTCCTTGGCATTGGCCAGTGCTCAGTCTGTGCTGACTCAG

*Eco*R I

VL (DPL-3)

3' end

230 235

A P T E C S O O

GCCCCTACAGAATGTTCATAATAA**GGCGCGCC**

Asc I

Cλ

B) *PCR amplification of the Fab light chain A2*

The Fab light chain gene segment was amplified by Expand™ PCR (Section 2.2.7c) using the vector, pUC119His6mycXba FITC-A2 as a DNA template and the primer pair CL(FORX)/Agg(L/L) [Table 3.6]. The reaction conditions were: [(94°C, 1min; 55°C, 2min; 72°C, 1min) x 25; 72°C, 10min] and the PCR product obtained was purified using the relevant DNA extraction kit (Section 2.2.9b).

C) *Construction of the vector, pYX-Light A2*

The expression vector, pYX243 was digested with *Eco*R I (20units) and *Nhe* I (5units) overnight, at 37°C (Section 2.2.12). The digested pYX243-*Eco*R I/*Nhe* I vector (8.5k.b.p.) was resolved by low melting agarose gel electrophoresis (Section 2.2.8) and purified (Section 2.2.9b). The Fab light A2 PCR product was treated with restriction enzymes, *Eco*R I (20units) and *Xba* I (20units) overnight, at 37°C (Section 2.2.12) and the digestion buffers were removed (Section 2.2.9b).

The Fab light A2-*Eco*R I/*Xba* I insert was ligated into vector, pYX243-*Eco*R I/*Nhe* I (Section 2.2.13) and transformed in *E. coli* (Section 2.2.15). Ampicillin-resistant colonies were PCR screened (Section 2.2.16) using primers pYX(FOR2) and pYX(BACK2). Positive transformants were grown overnight in 3ml of LB (Section 2.2.1) and 'minipreps' were prepared (Section 2.2.17). Diagnostic restriction digestions with enzymes *Eco*R I (20units) and *Asc* I (10units) were performed on three positive colonies to check the light chain was present and glycerol stocks were made (Section 2.2.3). Primers pYX(FOR2), PYX(BACK2) and CI.lib.seq were used to sequence the light chain (Table 3.6).

Table 3.6 – The sequences of primers used in the construction and verification of the light chain. Restriction endonuclease sites are in bold.

Primer Name	Sequence (5'-3')	Description (5'-3')
CL(FORX)	GCT**CTAGAGGCGCGCC**TTATTATGAACA TTCTGT	2 n.t. overhang; a *Xba* I site; an *Asc* I site; 18 n.t. of homology to the 3' end of human λ constant region (non-coding strand).
Agg(L/L)	CCG**GAATTC**ATGTTCACTTTTCTCAAAAT TATTCTGTGGCTTTTTTCCTTGGCATTGG CTAGCGCTCAGTCTGTGCTGACTCAG	3 n.t. overhang; an *Eco*R I site; 57 n.t. α-agglutinin signal sequence; 18 n.t. of homology to the 5' end of the human variable light chain, DPL-3 (coding strand).
PYX(FOR2)	TGTATTACTTCTTATTCA	Binds 76-94 b.p. upstream from the *Eco*R I site in pYX243 (coding strand).
PYX(BACK2)	TTAATACTAGGAACAAGA	Binds 35-53 b.p. downstream from *Nhe* I site in pYX243 (non-coding strand).
CI.lib.seq	GTGGCCTTGTTGGCTTGAAGC	Binds to the lambda constant region 440-461 b.p. downstream from the start codon of construct Fab A2 light (non-coding strand).

3.3.3 Transformation of *S. cerevisiae* cells with pYES2-*SAG1*-Heavy A2 and pYX-Light A2

Expression vectors, pYES2-*SAG1*-Heavy A2 (~2μg) and pYX-Light A2 (~2μg), were co-transformed (Section 2.2.22) into the *S. cerevisiae* strain, YMTAB. Leu^+ Ura^+ transformants were PCR screened (Section 2.2.16) for the presence of both plasmids in separate PCRs. The pYES2-*SAG1*-Heavy A2 expression vector was detected using primers pYX(FOR)/pYES(FOR2) and the pYX-Light A2 plasmid was identified using pYX(FOR2)/pYX(BACK2). Each PCR was analysed by agarose gel electrophoresis (Section 2.2.8) and positive transformants were grown in 1ml of SC lacking uracil and leucine (supplemented with 2% raffinose) for 24 hours at 30°C (Section 2.2.2).

Glycerol stocks were prepared (Section 2.2.3) and a 100µl aliquot of one culture was plated onto a SC agar plate containing 2% rafffinose but lacking uracil and leucine. This was used as a stock plate.

3.3.4 Construction of the scFv B7 expression vector

A) Primer design

The primer pairs, VH(B2)/VH(F) and VK(B)/VK(F/NT), were designed to amplify the human variable heavy (DP-67) and light (DPK-12) chains, respectively, from the vector pUC119His6mycXba-FITC B7 (Figure 3.10). The 5' ends of oligonucleotides VH(F) and VK(B) were engineered to be complementary to each other and encoded for the 15 residue $(Gly_4Ser)_3$ linker. Hence, in a second PCR, the amplified variable regions can be joined together via a 'splicing by overlap extension' reaction to create the single-chain Fv B7. To facilitate cloning into the expression vector, pYES2-*SAG1*-Heavy A2, the primers VH(B2) and VK(F/NT) were designed to incorporate recognition sites, *Msc* I and *Not* I, respectively. The engineered *Msc* I site in the pYES2-*SAG1*-Heavy A2 expression vector is present four nucleotides upstream from the 5' Heavy chain/signal sequence junction. Therefore, primer VH(B2) was also synthesized to encode for the last two amino acids of the signal sequence. Using this cloning strategy allowed the removal of the heavy chain A2 gene segment and the in-frame replacement of the scFv B7 gene. Figure 3.11 summarizes the scFv B7/α-agglutinin expression construct.

B) *PCR amplification of the light and heavy variable regions of Fab FITC-B7.*

The human heavy (DP-67) and light (DPK-12) variable regions were amplified from vector pUC119His6mycXba- Fab FITC B7 in separate Expand PCRs (Section 2.2.7.c) using primer pairs VH(B2)/VH(F) and VK(B)/VK(F/NT), respectively (Table 3.7). The reaction conditions for the amplification of the heavy chain were: [(94°C, 1min; 55°C, 1min; 72°C, 1min) x 25; 65°C, 5min] and the cycling parameters for the light chain were the same, except the annealing temperature was 60°C.

Both amplified fragments were resolved by low-melting point agarose gel electrophoresis (Section 2.2.8) and purified (Section 2.2.9b).

Figure 3.10 – Flow diagram summarizing the construction of the scFv B7 expression vector, pYES2-*SAG1*-scFv B7. The letters L, G, Xa and T represent the α-agglutinin leader sequence, $(Gly_4Ser)_3$ linker, the Factor Xa recognition site and a yeast terminator sequence, respectively. The symbols VL and VH depict the variable light gene segment (DPK-12) and the variable heavy gene segment (DP-67), respectively.

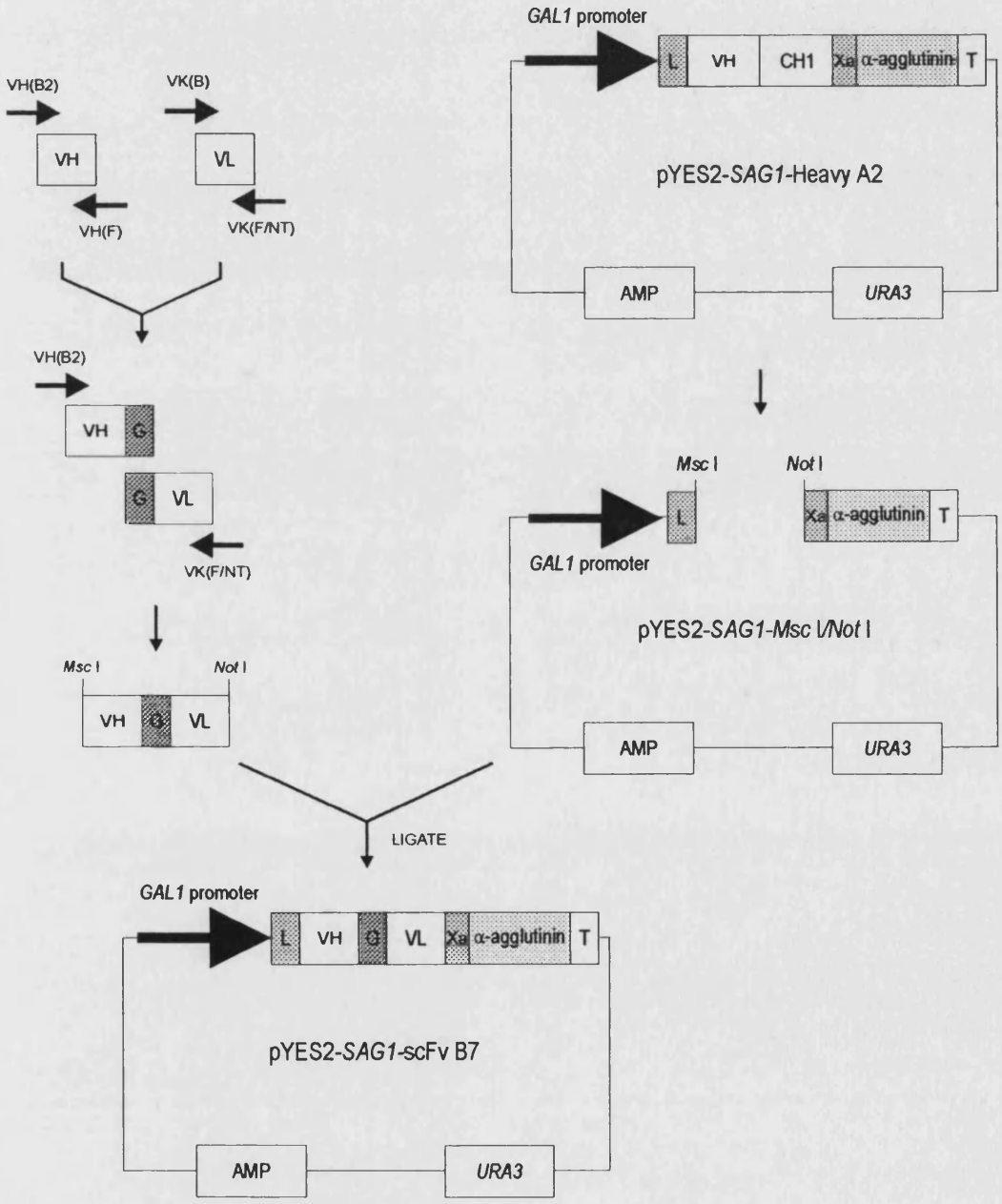
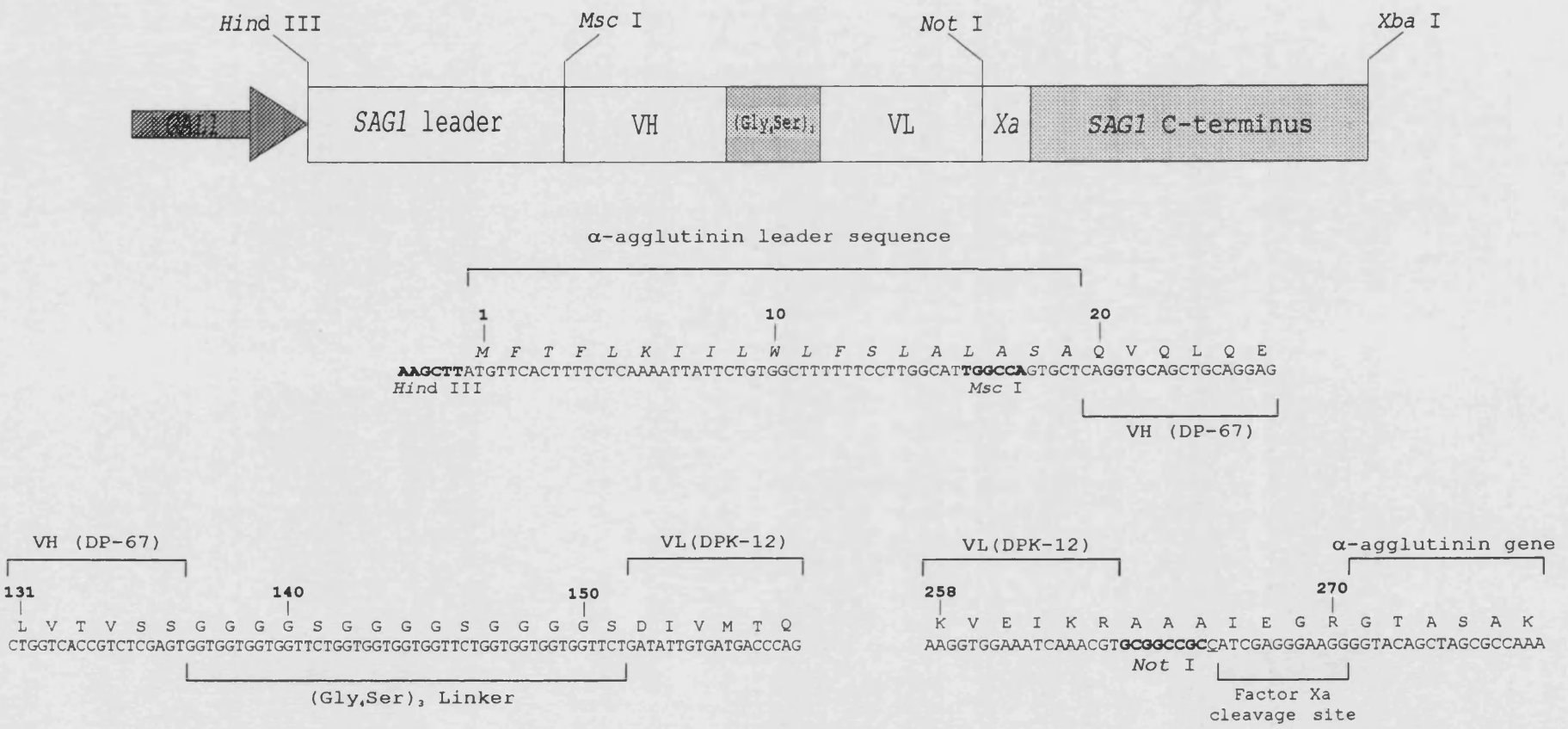


Figure 3.11 - Schematic representation of the single-chain Fv B7/α-agglutinin expression construct including sequence details of the 5' end. The DNA sequences encoding the $(Gly_4Ser)_3$ linker and scFv B7/α-agglutinin junction are also described. Restriction sites are represented in bold.



C) *PCR assembly of the single-chain Fv B7*

Approximately equal amounts (~50ng) of the light (DPK-12) and heavy (DP-67) variable regions were added together in a $Vent_R$ PCR (Section 2.2.7a) with primer pair VH(B2)/VK(F/NT). The reaction was performed for 25 cycles using the following parameters: [(94°C, 1min; 65°C, 2min; 72°C, 1min) x 25; 72°C, 10min]. The assembled product (767 b.p.) was separated from the other non-specific bands by low-melting point agarose gel electrophoresis (Section 2.2.8) and recovered using β-agarase (Section 2.2.9a).

D) *Construction of expression vector, pYES2-SAG1-scFv B7*

The expression vector, pYES2-*SAG1*-Heavy A2 and the amplified scFv B7 fragment were digested with *Msc* I (15units) and *Not* I (10units) overnight, at 37°C (Section 2.2.12). The pYES2-*SAG1-Msc I/Not* I was separated from the heavy A2 chain gene segment by low melting point agarose gel electrophoresis (Section 2.2.8) and purified (Section 2.2.9b). Additionally, the restriction buffers were removed from scFv B7-*Msc I/Not* I using the appropriate DNA extraction kit (Section 2.2.9b). ScFv B7-*Msc I/Not* I was ligated into pYES2-*SAG1-Msc I/Not* I (Section 2.2.13) and transformed into *E. coli* cells (Section 2.2.15). The ampicillin-resistant colonies were PCR screened (Section 2.2.16) using primers pYES(FOR2)/pYX(FOR) and conditions: [(94°C, 1min; 50°C, 1min; 72°C, 2min) x 25; 72°C, 10min]. Positive transformants were grown overnight in 3ml of LB (Section 2.2.1) and plasmid DNA was prepared from each culture (Section 2.2.17). Diagnostic restriction digests with *Xba* I (20units) and *Hind* III (10units) were performed and glycerol stocks were made (Section 2.2.3). DNA sequencing (Section 2.2.18) was performed using primers pYX(FOR) and AggSeq1 (Table 3.7).

3.3.5 Transformation of *S. cerevisiae* cells with pYES2-*SAG1*-scFv B7

The expression vector, pYES2-*SAG1*-scFv B7 was transformed into the protease-deficient *S. cerevisiae* strain, YMTAB (Section 2.2.22). Ura^+ colonies were PCR screened with primers pYES2(FOR2) and pYX(FOR), using conditions: [(94°C, 1min; 50°C, 1min; 72°C, 2min) x 25; 72°C, 10min].

Positive transformants were grown in 1ml of SC containing 2% raffinose but lacking uracil for 24 hours (30°C; Section 2.2.2) Glycerol stocks were prepared (Section 2.2.3) and a 100µl aliquot was placed onto a SC without uracil/2%(w/v) raffinose agar plate (Section 2.2.2). Additionally, the pYES2 expression vector was transformed into YMTAB for a negative expression control.

Table 3.7 – Sequences of the oligonucleotides used in the construction of the scFv B7 expression vector. Enzyme recognition sites are in bold and the sequence encoding the $(Gly_4Ser)_3$ linker is underlined .

Primer Name	Sequence (5'-3')	Description (5'-3')
VH(B2)	CCCCCC**TGGCCA**GTGCTCAGGTGCAGC TGCAGGAG	6 n.t. overhang; a *Msc* I site; the last four n.t. of the α-agglutinin leader sequence; 18 n.t. of homology to the 5' end of germline sequence, DP-67 (coding strand)
VH(F)	AGAACCACCACCACCAGAACCACCACCA CCAGAACCACCACCACCACTCGAGACG GTGACCAG	45 n.t. encoding the $(Gly_4Ser)_3$ linker; 18 n.t. of homology to the 3' end of heavy variable sequence DP-67 (non-coding strand).
VK(B)	GGTGGTGGTGGTTCTGGTGGTGGTGGT TCTGGTGGTGGTGGTTCTGATATTGTGA TGACCCAG	45 n.t. encoding the $(Gly_4Ser)_3$ linker; 18 n.t. of homology to the 5' end of the variable kappa germline sequence, DPK-12 (coding strand).
VK(F/NT)	AAGGAAAAAA**GCGGCCGC**ACGTTTGATT TCCAG/CCTTGGT	10 n.t. overhang; a *Not* I site; 21 n.t. to the 3' end of the variable kappa germline sequence, DPK12 (non-coding strand).

3.4 RESULTS

3.4.1 Construction of the expression vector, pYES2-*SAG1*-Heavy A2

The construction of the heavy chain A2/α-agglutinin gene fusion was successfully achieved in two PCR steps.

In the first step, the anti-FITC Fab heavy chain A2, which consisted of the human variable germline sequence, DP-47 and the first human constant domain of the IgG1 molecule, was amplified from the *E. coli* expression vector, pUC119His6mycXba-FITC A2 (Griffiths *et al.,* 1994). Additionally, the C-terminal anchorage domain of the α-agglutinin gene was amplified from plasmid pH27 (Lipke *et al.*, 1989). The sizes of both the generated PCR fragments were determined by agarose gel electrophoresis (Figure 3.12) and were in agreement with the predicted theoretical sizes i.e. 749 b.p. for the heavy chain A2 and 986 b.p. for the α-agglutinin anchorage domain.

In the second PCR, these two fragments were linked together by a 'splicing by overlap extension' reaction and the heavy chain A2/α-agglutinin gene fusion was created (Figure 3.12). The expression vector, pYES2 and the gene fusion, heavy chain A2/α-agglutinin, were digested with *Hin*d III and *Xba* I and ligated together. Transformation of the construct produced several hundred plaques and ten of these were PCR screened for the heavy/α-agglutinin insert.

Eight transformants were found to be positive and DNA sequencing of this final pYES2-*SAG1*-Heavy A2 vector confirmed an in-frame fusion between the α-agglutinin signal sequence, heavy chain A2 gene segment and the 3' half of the α-agglutinin gene had been created (Figure 3.7). Three heavy chain expression vectors were sequenced and all contained a nucleotide change in the first IgG1 constant domain. This nucleotide change converted an asparagine amino acid to a lysine residue at amino acid position 225 of the heavy chain A2 /α-agglutinin fusion (Figure 3.13) and was considered not to be an error that occurred during the polymerase chain reaction because it was present in all of the constructs.

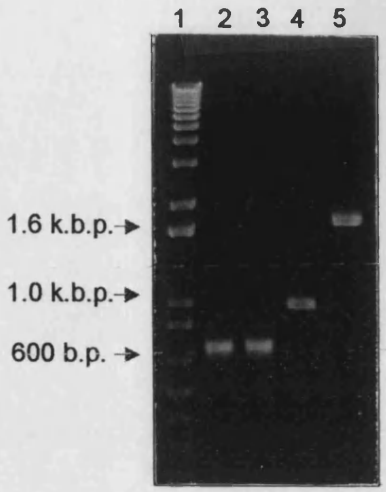


Figure 3.12 - A 1% agarose gel showing the various fragments used in the construction of the expression vectors, pYES2-*SAG1*-Heavy A2 and pYX-Light A2. Lane 1, 1 k.b.p. DNA marker; Lane 2, Fab light chain A2 (736 b.p.); Lane 3, Fab heavy chain A2 (749 b.p.); Lane 4, α-agglutinin anchorage domain (986 b.p.); Lane 5, Heavy chain A2-*SAG1* gene fusion (1716 b.p.).

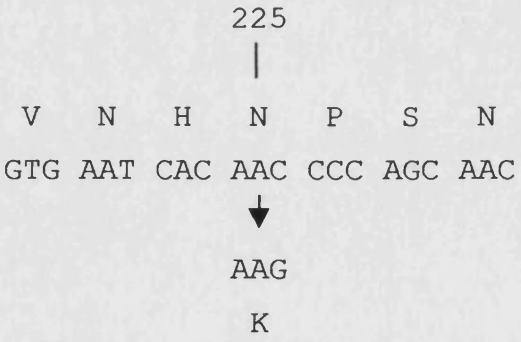


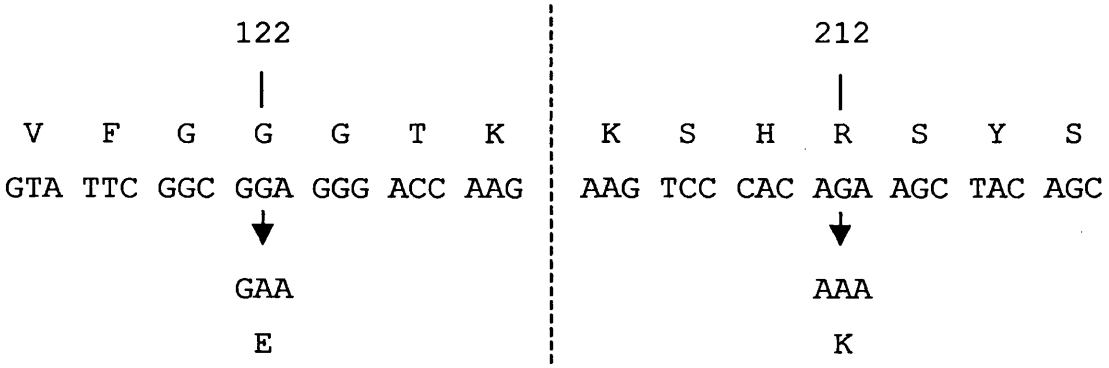
Figure 3.13 – The nucleotide change (amino acid position 225) found in the IgG1 constant region of the heavy chain A2/α-agglutinin gene construct when compared to the published human IgG1 constant sequence (Kabat *et al.,* 1991). The numbering of the amino acids refers to the complete heavy A2/α-agglutinin fusion as indicated in Figure 3.7.

3.4.2 Construction of the light chain expression vector, pYX-Light A2

The anti-FITC Fab light chain A2 which consisted of the human variable germline sequence, DPL-3, and the lambda constant domain was amplified from vector, pUC119His6mycXba-FITC A2. A single PCR product of

approximately 750 b.p. was obtained (Figure 3.12) and digested with enzymes, *Eco*R I and *Xba* I. The *Eco*R I/*Xba* I digested light chain was cloned into the yeast expression vector, pYX243-*Eco*R I/*Nhe* I due to the compatible cohesive ends of restriction sites, *Nhe* I and *Xba* I. Transformation of the construct produced several hundred plaques and ten of these were PCR screened for the light chain insert. Four were found to be positive and this result was confirmed by digestion with enzymes, *Asc* I and *Eco*R I. DNA sequencing of the construct revealed the α-agglutinin signal sequence was attached correctly to the light chain gene segment (Figure 3.9). Two nucleotides in the light chain A2 construct differed from the published sequence at amino acid positions 122 (G122E) and 212 [R212K (Figure 3.14)]. These were presumed not to be errors that occurred during the amplification process due to the fact they were present in all of the constructs sequenced.

Figure 3.14 – The nucleotide changes (amino acid positions 122 and 212) discovered during DNA sequencing of the light chain A2 gene fusion. Numbering of the amino acids refers to the total light A2 construct as shown in Figure 3.9.



3.4.3 Co-transformation of pYX-light A2 and pYES2-*SAG1*-Heavy A2

The anti-FITC Fab expression vectors, pYX-light A2 and pYES2-*SAG1* Heavy A2, were co-transformed into the protease-deficient *S. cerevisiae* strain, YMTAB. Approximately two hundred colonies were obtained for this transformation and ten were PCR screened to check both plasmids were present.

3.4.4 Construction of the pYES2-*SAG1*-scFv B7 expression vector

The human heavy (DP-67) and light (DPK-12) variable regions were amplified from vector, pUC119His6mycXba-FITC B7. The sizes of both the generated PCR fragments were determined by agarose gel electrophoresis (Figure 3.15) and were in good agreement with the predicted theoretical sizes i.e. 413 b.p. for the human variable heavy chain (DP-67) and 399 b.p. for the human variable light chain (DPK-12).

The 3' end of the variable heavy chain and the 5' end of the variable light chain were complementary to each other and encoded a $(Gly_4Ser)_3$ linker. Therefore, in a second PCR step, the heavy and light chains were joined together by a 'splicing by overlap extension' reaction to create the single-chain Fv B7 (Figure 3.15).

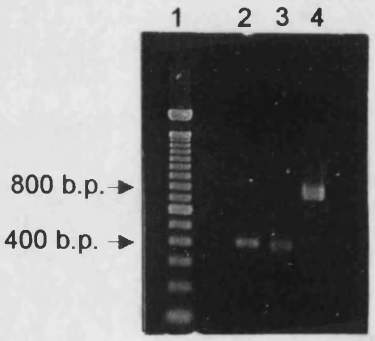


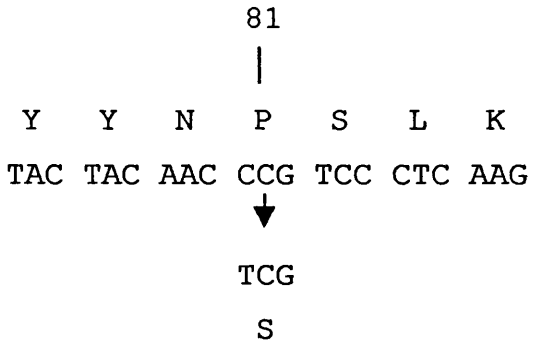
Figure 3.15 – A 1.5% agarose gel showing the construction of the single chain Fv B7. Lane 1, 100b.p. DNA marker; Lane 2, variable heavy chain (413 b.p.); Lane 3, variable light chain (399 b.p.); Lane 4, single chain Fv (767 b.p.).

Digestion of the pYES2-*SAG1*-Heavy A2 expression vector with *Msc* I and *Not* I removed the heavy chain gene segment. The digested scFv–*Msc* I/*Not* I fragment was then inserted into this digested expression vector to create the pYES2-*SAG1*-scFv B7 construct. This was transformed into *E. coli* and twelve ampicillin-resistant colonies were PCR screened. All twelve colonies contained the scFv B7 insert and DNA sequencing confirmed the α-agglutinin leader sequence and anchorage domain flanked the anti-FITC scFv B7 gene (Figure 3.11). One nucleotide, at amino acid position 81

(Figure 3.16), differed from the published sequence in the variable heavy region (DP-67) in all five of the constructs sequenced. This change converted a proline to a serine amino acid (P81S) in the CDR2.

The pYES2-SAG1-scFv B7 expression vector was transformed into the protease-deficient *S. cerevisiae* strain, YMTAB. Ten colonies were screened from this transformation to confirm the plasmid was present.

Figure 3.16 – The nucleotide change (amino acid position) found in the variable heavy chain sequence (DP-67) of the scFv B7 gene. Amino acid numbering refers to the scFv B7/α-agglutinin construct as shown in Figure 3.11.



3.5 DISCUSSION

Both the heavy and light chain surface display vectors, pYES2-*SAG1*-Heavy A2 and pYX-Light A2, respectively, were successfully constructed for the expression of the anti-fluorescein-5-isothiocyanate (FITC) Fab antibody fragment on the surface of *S. cerevisiae.*

The heavy chain surface display vector was constructed by separately amplifying the heavy chain gene segment and the 3' half of the α-agglutinin gene and fusing these two fragments together using a 'splicing by overlap extension' reaction (Figure 3.6). An α-agglutinin signal sequence was attached to the 5' terminus of both the heavy (DP-47 gene segment + first human IgG constant heavy region) and light (DPL-3 + human lambda constant region) chains, to target expression to the secretory pathway (Figure 1.9; Section 1.6.2). The light and heavy chain constructs were cloned into separate expression vectors (Sections 3.4.1 and 3.4.2) containing the

same inducible promoter, *GAL1*. The two expression vectors used, pYX243 and pYES2, contained different yeast auxotrophic markers, therefore were co-transformed in the relevant *S. cerevisiae* strain for expression. The *S. cerevisiae* strain used was deficient in two of the major vacuolar proteases (prA and prB). The proteolysis problem associated with cell breakage in yeast (Jones, 1991) should be overcome using this strain and theoretically will help in the protein analysis of cell wall extracts.

It was presumed, after induction with galactose, that the heavy and light chains would be simultaneously expressed through the secretory pathway and form functional Fab fragments in the endoplasmic reticulum. Hopefully, the Fab antibody fragment would then be transported through the rest of the secretory pathway to the cell surface where the α-agglutinin anchorage domain, which is attached to the C-terminus of the heavy chain, would become covalently associated to the β-glucan component of the cell wall. This attachment should lead to the surface display of the anti-FITC Fab fragments. Functional Fab fragments have been secreted into the yeast culture supernatant after simultaneously expression of the light and heavy chains from different expression vectors, therefore this strategy should be plausible (Horwitz *et al.*, 1988). Interestingly, the display of Fab fragments on the surface of phage utilizes the minor phage coat protein in a similar way (Section 1.4.4; Hoogenboom *et al.*, 1991).

A Factor Xa cleavage site was placed between the heavy chain gene segment and the α-agglutinin anchorage domain to assist in detachment of the Fab molecules from the yeast cell surface. Detachment may ease the characterization and purification of the antibody fragments. This part of the strategy is intended to mimic the soluble antibody producing properties of plasma B-cells.

The successful construction of the anti-FITC scFv surface display construct, pYES2-*SAG1*-scFv B7, was aided by the careful design of the heavy chain expression vector. It was speculated that expression of the functional scFv antibody fragment on the surface of yeast might be more successful than the Fab fragment. The reasoning for this speculation was that the variable heavy and light chains of the scFv molecule would be

expressed in eqimolar concentrations and would not rely on non-covalent association of the individually expressed light and heavy chains in the endoplasmic reticulum to assemble a functional antibody fragment.

The joining of the amplified variable heavy (DP-67) and light (DPK-12) chains using a 'splicing by overlap extension' reaction created the anti-FITC scFv B7 antibody fragment (Section 3.4.4). The complementary oligonucleotides used in this 'splicing by overlap extension' reaction encoded for a $(Gly_4Ser)_3$ linker (Huston *et al.*, 1988; Section 1.4.3). This linker should be sufficiently long and flexible enough to span the distance between the variable domains and allow the formation of the antigen-combining site. In this project, the linker fused the C-terminus of the variable heavy chain to the N-terminus of the variable light chain. This orientation should not affect the binding constant of the single chain Fv for the fluorescent antigen, FITC (Bird and Walker, 1991) and hopefully an affinity comparable with the anti-FITC Fab B7 ($Ka = 6.6 \times 10^6\ M^{-1}$) should be achieved.

In the heavy chain construct, two unique restriction sites were designed at the 3' ends of the α-agglutinin signal sequence and heavy chain. These restriction sites were used to replace the heavy chain with the single chain Fv construct. Therefore, the created anti-FITC scFv B7 gene was flanked by the α-agglutinin signal sequence and 3' half of the α-agglutinin gene, downstream from the inducible *GAL1* promoter. The engineered pYES2-*SAG1*-scFv B7 expression vector contains the correct molecular information to over-express the scFv antibody fragments on the surface of *S. cerevisiae.*

DNA sequencing of all three constructs revealed the germline segments differed by one to two nucleotides from the published sequences (Kabat *et al.*, 1991). It was presumed that these nucleotide changes were not errors that had occurred during the amplification process because they were present in all constructs sequenced. Therefore, these nucleotide changes must have been present in the supplied *E. coli* expression vectors, pUC119His6mycXba-FITC B7 and pUC119His6mycXba-FITC A2. The reasons why these changes were present could be numerous. They may have been introduced in previous polymerase chain reaction experiments

(Griffiths *et al.*, 1994). Alternatively, they may be due to polymorphisms or allelic variations between unrelated individuals. In the next chapter, both the anti-FITC Fab constructs (A2 and B7) were expressed in *E. coli* to check if they had retained their affinity for the fluorescent antigen, FITC (Section 4.4.1).

Conclusions

The aims of this Chapter were achieved, mainly due to the careful design of the oligonucleotides used in the construction of the three expression vectors. Both the pYES2-*SAG1*-Heavy A2 and the pYX-Light A2 expression vectors were constructed for the over-expression of the anti-FITC Fab antibody fragment on the surface of yeast. Additionally the display vector, pYES2-*SAG1*-scFv B7, was assembled for surface expression of the anti-FITC scFv antibody fragment. Expression of these constructs on the surface of *S. cerevisiae* is described in Chapter 4.

CHAPTER 4

EXPRESSION OF ANTIBODY FRAGMENTS ON THE SURFACE OF YEAST

4.1 INTRODUCTION

The preceding chapter described the construction of the expression vectors required for the display of the anti-FITC Fab A2 antibody fragment on the surface of *S. cerevisiae.* Additionally, the anti-FITC single chain Fv expression vector, pYES2-*SAG1*-scFv B7, was created. These constructs comprised of the appropriate antibody DNA flanked by the α-agglutinin signal sequence and the C-terminal half of the α-agglutinin gene, all under the control of the inducible *GAL1* promoter (Figure 3.3). After induction with galactose, it is presumed the α-agglutinin leader sequence will target expression to the secretory pathway (Section 1.6.2) and the C-terminal α-agglutinin anchorage domain will covalently attach the antibody fragment to the glucan component of the cell wall. (Section 1.7). The 5' binding domain of the α-agglutinin gene is replaced by the antibody DNA, therefore once expressed, the antibody fragments should be exposed on the exterior surface of the cell wall.

This chapter investigates the individual expression of the anti-FITC Fab and scFv antibody fragments on the surface of *S. cerevisiae.* Schreuder and co-workers (1993) showed that non-covalently attached proteins were released by incubation with hot sodium dodecyl sulphate (SDS). After SDS treatment, proteins attached to the glucan component of the cell wall can be liberated by enzymatic digestion with glucanase. This approach was recruited in this project to determine if the antibody-agglutinin protein fusions were covalently attached to the cell wall's glucan.

Expression of the antibody fragments on the surface of *S. cerevisiae* cells was investigated using antibody detection methods such as enzyme

immunoassays (Wojciechowicz and Lipke, 1989) and indirect immunofluoresence (Schreuder *et al.*, 1993). Antibodies cannot penetrate the yeast cell wall (Pringle *et al.*, 1991), therefore these experiments should determine if the antibody fragments are on the exterior surface of *S. cerevisiae* and are accessible to macromolecules. Finally, the functionality of the antibody fragments on the surface of *S. cerevisiae* was examined using the fluorescence antigen, fluorescein-5-isothiocyanate (FITC).

Glycosylation can affect both the molecular weight and detection levels of the antibody-agglutinin fusions therefore the glycosylation process in yeast is reviewed below:

4.1.1 The glycosylation process

Proteins can be modified in various ways during their passage through the secretory pathway (Section 1.6.2). The addition of carbohydrates to the protein backbone (glycosylation) first begins in the endoplasmic reticulum (Romanos *et al.*, 1992). A core oligosaccharide unit comprising of two N-acetyl-glucosamine, nine mannose and three glucose residues are added to asparagine residues (N-glycosylation) in the Asn-Xxx-Ser/Thr consensus sequence. The glucose residues are subsequently trimmed from the unit and one mannose residue is removed. This core N-glycosylation step is common to yeast, plants and higher eukaryotes.

Modification of the core oligosaccharide unit takes place in the Golgi apparatus and at this stage there is divergence between the glycosylation process in yeast and other eukaryotes. In *S. cerevisiae,* the mannose residues are not trimmed back as in higher eukaryotes and instead up to 75 residues can be added to the core oligosaccharide unit. This process is known as hyperglycosylation and gives rise to very heterogenous proteins with varying degrees of oligosaccharide attachment.

Additionally, oiligosaccharides can be O-linked to serine and threonine residues in the endoplasmic reticulum (O-glycosylation). Extension of O-linked chains up to a final length of five mannose residues takes place in the Golgi. Despite extensive research, no clear rules have emerged about how

the cells select the serine and threonine residues for O- glycosylation (Orlean *et al.*, 1991).

The C-terminal α-agglutinin anchorage domain is very rich in serine and threonine residues and contains seven potential N-glycosylation sites (Hauser and Tanner, 1989). Both N- and O-linked oligosaccharides are attached to the mature cell wall form of the agglutinin molecule (Lu *et al.*, 1994). Hence, glycosylation of the antibody-agglutinin protein fusions will affect their molecular sizes and very heterogenous proteins may be observed.

4.2 MATERIALS

4.2.1 Reagents

The protease inhibitors antipain, chymostatin, and pepstatin were purchased from Sigma. Additionally, lyticase, *o*-nitrophenyl-β-D-galactopyranoside (ONPG), raffinose, galactose, albumin bovine-fluorescein-5-isothiocyanate (FITC-BSA) and 0.2μ glass beads were all obtained from Sigma (Product numbers – L5763, N1127, R0514, G0750, A9771 and G1145, respectively). Phenylmethylsulphonyl fluoride (PMSF) and leupeptin were supplied by Boehringer Mannheim. Factor Xa protease was purchased from New England Biolabs and mounting medium was obtained from Cityfluor (Code No. - R1320). The concentrators, Centriprep® 10 and Microcon® 10, were purchased from Amicon, Beverly, USA.

4.3 METHODS

4.3.1 Expression and functional detection of anti-FITC A2 and B7 Fab fragments in *E. coli* (EMBO practical course, 1991)

Two individual TG1 *E. coli* colonies carrying either expression vector pUC119His6mycXba-FITC A2 or pUC119His6mycXba-FITC B7 were picked from an LB plate (supplemented with 100μg/ml ampicillin). The colonies were resuspended in wells (using a 96-well plate, Nuclon) containing 200μl of XL-

media (16g/l Bacto-tryptone, 16g/l Bacto-yeast extract, 5g/l soduim chloride), 100 μg/ml ampicillin and 1% (w/v) glucose. These cultures were grown overnight (300 r.p.m.) at 37°C and 10μl of each overnight culture was transferred to a second 96-well plate containing 200μl of XL-media, 0.1% (w/v) glucose and 100μg/ml ampicillin, per well. The new cultures were grown at 37°C until an absorbance (at 600nm) of 0.9 was reached. Twenty-five microlitres of XL-media supplemented with 9mM isopropyl-thiogalactoside (IPTG) and 100μg/ml ampicillin was added to each well and incubated at 30°C for an additional 24 hours. After this time, the cultures were centifuged at 3000 r.p.m. (1500g) for 10 minutes using a GLC4 general laboratory centifuge. Each of the culture supernatants (100μl) was analysed for functional anti-FITC Fab fragments using an ELISA (Section 2.3.4). The supernatant from a TG1 wild type strain culture was used throughout the ELISA as a negative control.

A solution of albumin bovine-fluorescein-5-isothiocyanate (FITC-BSA) at a concentration of 1.5×10^{-6} M was used to coat the wells of the ELISA plate. Three different sets of antibodies were individually used to analyse the culture supernatant. Firstly, a primary polyclonal antibody serum specific for human IgG1 Fab fragments (goat anti-human IgG Fab; Table 2.1) was used to analyse both the Fab A2 and B7 culture supernatants. A complementary horseradish-peroxidase conjugated rabbit anti-goat IgG secondary antibody (Table 2.2) was used for detection.

The second type of antibody used was specific for human kappa light chains (rabbit anti-human kappa light chain; Table 2.1) and was utilised to detect the Fab B7 fragment that contains the human variable (DPK-12) and constant kappa light chain. Binding of this primary antibody was detected using a horseradish peroxidase conjugated goat anti-rabbit IgG secondary antibody (Table 2.2). The third type of antibody was specific for human lambda light chains (mouse anti-human lambda light; Table 2.1) and was recruited to detect the Fab A2 antibody fragment that comprises of a human lambda variable (DPL-3) and constant light chain. This primary antibody was detected using the secondary horseradish peroxidase conjugated goat anti-mouse IgG antibody (Table 2.2)

Periplasmic lysate preparation

Periplasmic lysates from *E. coli* cells expressing either the anti-FITC Fab A2 or B7 antibody fragments were prepared for positive controls in Western blotting analysis (Section 2.3.3).

Two individual TG1 *E. coli* colonies, carrying either expression vector pUC119His6mycXba-FITC A2 or pUC119His6mycXba-FiTC B7 were picked from a LB plate (supplemented with 100μg/ml ampicillin). These colonies were grown overnight at 37°C in 50ml of XL-media supplemented with 100μg/ml ampicillin and 0.1% (w/v) glucose. After this time, isopropyl-thiogalactosidase (IPTG) was added to a final concentration of 1mM and the cultures were incubated at 26°C for a further 3 hours. The cells were then collected by centrifugation (4°C) at 8000g for 10 minutes using an IRC-5B refrigerated superspeed centrifuge (Sorvall). The cells were resuspended in 10ml of lysis buffer [30mM Tris-HCl, pH 8.0, 1mM EDTA, 20% (w/v) sucrose] and shaken for 10 minutes at room temperature. The centrifugation step was repeated and the cells were resuspended in 10ml of 5mM magnesium sulphate. This cell mixture was agitated at 4°C for 10 minutes and the centrifugation step was repeated. The supernatant (periplasmic lysate) was decanted and stored at -20°C until required.

4.3.2 Optimising galactose induction time using a β-galactosidase assay (Adams *et al.*, 1996)

The supplied control plasmid, pYX243-lacZ (R+D systems) was transformed into *S. cerevisiae* strain, YMTAB (Section 2.2.22). Leu^+ colonies were PCR screened (Section 2.2.16) using primers pYX(FOR2) and pYX(BACK2) and the PCR products were analysed by agarose gel electrophoresis (Section 2.2.8).

A single positive colony was transferred from a SC agar plate lacking leucine [supplemented with 2% (w/v) raffinose] into one millilitre of SC medium without leucine/2% (w/v) raffinose. The culture was grown for 24 hours (250 r.p.m.) at 30°C, counted using a haemocytometer and diluted in the same medium to 3-5 x 10^6 cells/ml. Galactose was added to a final concentration of 2% (w/v) and the culture was incubated (250 r.p.m.) at 30°C

Approximately 1 x 10^6 to 1 x 10^7 cells were taken from the culture directly after galactose induction and every 1.5-2 hours thereafter. An absorbance reading (at 600nm) was recorded for each culture fraction directly after sampling and the cells were harvested (4°C) at 2000 r.p.m. for 5 minutes. The cells were resuspended in 1ml of ice-cold Z buffer, pH 7.0 [58mM $Na_2HPO_4.7H_2O$, 38mM $NaH_2PO_4.H_2O$, 10mM KCl, 1mM $MgSO_4.7H_2O$, 0.27% (w/v) β-mercaptoethanol]. Three drops of chloroform and two drops of 0.1% (w/v) sodium dodecyl sulphate (SDS) were added and the mixture was vortexed vigorously for 10 seconds. The sample was then incubated at 28°C for 5 minutes. A 4mg/ml solution of *o*-nitrophenyl-β-D-galactoside (ONPG) in ice-cold Z buffer was prepared and 0.2ml of this ONPG solution was added to each sample. When a pale yellow colour developed, the reaction was stopped with 0.5ml of 1M sodium carbonate. The time of the reaction was measured from the addition of the substrate, ONPG, to termination with sodium carbonate. An absorbance (at 420nm) was measured for each sample tested and β-galactosidase activity was calculated using the following equation.

$$\beta\text{-galactosidase units} = \frac{Abs_{420}}{Abs_{600} \times \text{volume} \times \text{time}}$$

Abs_{420} = absorbance (at 420nm) of the reaction product, o-nitrophenol.
Abs_{600} = absorbance (at 600nm) of the culture at the time of the assay.
Volume = amount of the culture used in the assay.
Time = time of the reaction in minutes.

4.3.3 Galactose induction of antibody fragments onto the surface of *S. cerevisiae*

All surface expression experiments in this chapter were induced with galactose using the following method:

A single colony containing the relevant expression vectors was grown for 24 hours at 30°C in 1ml of selective medium lacking the appropriate amino acids for maintenance of the plasmids[a] and supplemented with 2%

(w/v) raffinose. The cells were counted using a haemocytometer and diluted to 5 x 10^6 cells/ml (early exponential phase) in the correct selective medium containing 2% (w/v) raffinose. Galactose was added to a final concentration of 2% (w/v) and the cultures were grown for 24 hours at 30°C.

[a] = *S. cerevisiae* strain, YMTAB containing the anti-FITC Fab expression vectors, pYX-Light A2 and pYES2-*SAG1*-Heavy A2, was grown in SC media lacking uracil and leucine. The YMTAB strain carrying the anti-FITC scFv surface display construct, pYES2-*SAG1*-scFv B7, was grown in SC media without uracil.

4.3.4 Isolation and SDS-extraction of cell walls (Schreuder *et al.*, 1993)

A) Cell wall isolation

After galactose induction, the cells were collected for 2 minutes at 3000 r.p.m. (1500g) using a Sorvall GLC-4 general laboratory centrifuge. The cells were washed three times in ice-cold isolation buffer (10mM Tris-HCl, pH7.8, 1mM PMSF) and all subsequent steps were performed at 4°C. Three millilitres of isolation buffer and 10g of acid-washed acid beads (diameter - 0.2μM) were added to one gram of cells (fresh weight). The mixture was shaken vigorously on a horizontally rotating platform (ELISA plate shaker, Labline instruments Inc.) at 500 r.p.m. for 30 minutes. The supernatant was isolated and the glass beads were washed with 1M NaCl and 1mM PMSF, until the washes were clear. The supernatant and the glass bead washes were pooled and the cell walls were recovered by centrifugation for 5 minutes at 2445 r.p.m. (1000g) using a Sorvall GLC-4 general laboratory centifuge. The cell walls were then washed three times with 1mM PMSF and stored at -20°C.

B) SDS-extraction

Non-covalently bound proteins or proteins bound through disulphide bridges can be released from the cell walls by boiling in sodium dodecyl sulphate (SDS). The isolated cell walls were weighed into a 1.5ml Eppendorf tube and the weight was recorded. SDS-extraction buffer [50mM Tris-HCl, pH 7.8, 2% (w/v) SDS, 100mM EDTA, 40mM β-mercaptoethanol] was added to

the cell walls at a ratio of 2 : 1 (SDS-extraction buffer : cell walls). The cell walls were boiled for 5 minutes at 100°C and the supernatant containing the SDS-extractable material was collected by centrifugation at 13000 r.p.m. for 5 minutes. The supernatant (20µl) was resolved by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE; Section 2.3.1) and analysed by Western blotting (Section 2.3.3) using the appropriate sets of antibodies.
Note – This procedure was repeated until all the SDS-extractable material had been removed from the cell walls.

4.3.5 Identification of the individual light and heavy Fab chains in the SDS–extractable supernatant

After galactose induction (Section 4.3.3), cell walls were isolated (Section 4.3.4a) from an *S. cerevisiae* culture theoretically expressing the anti-FITC Fab fragments on its surfaces. The isolated cell wall fraction was treated with hot SDS (Section 4.3.4b) and analysed for the individual light and heavy Fab chains by Western blotting (Section 2.3.3). Three sets of antibodies were used:

1) A polyclonal antibody serum that was specific for all human IgG1 Fab fragments (Goat anti-human Fab IgG; Table 2.1) followed by a horseradish peroxidase conjugated rabbit anti-goat IgG secondary antibody (Table 2.2) for detection.
2) A monoclonal primary antibody, which was specific for human lambda light chains (Mouse anti-human lambda light specific; Table 2.1), and a complementary secondary horseradish peroxidase conjugated goat anti mouse IgG antibody (Table 2.2) for detection.
3) A primary antibody specific for human heavy chains (rabbit anti-human gamma chain specific; Table 2.1) followed by the secondary peroxidase conjugated goat anti-rabbit IgG antibody (Table 2.2) for detection.

All western blots were run with the SDS-extracted material from a wild type YMTAB culture for a negative control.

4.3.6 Detection of the individual light and heavy Fab chains on the surface of *S. cerevisiae*

The *S. cerevisiae* strain, YMTAB carrying the anti-FITC Fab A2 expression vectors, pYX-Light A2 and pYES2-*SAG1*-Heavy A2, was induced for 24 hours with galactose (Section 4.3.3) The heavy and light chains were individually detected on the surface of yeast using two different enzyme immunoassays (Section 2.3.5). The Fab A2 light chain was detected on the surface using an antibody specific for human lambda light chains. An antibody specific for human heavy chains was recruited to detect for surface display of the Fab A2 heavy chain. The details of both sets of antibodies used are described in Section 4.3.5. Each enzyme immunoassay was performed three times on the same induced culture and the wild type strain (YMTAB) was used as a negative control.

4.3.7 Confirming covalent attachment of the scFv B7 to the β-glucan component of the cell wall

The *S. cerevisiae* strain, YMTAB harbouring the anti-FITC scFv B7 expression vector, pYES2-*SAG1*-scFv B7 was induced with galactose for 24 hours (Section 4.3.3). After this time, the cell walls were isolated and treated with SDS-extraction buffer (Section 4.3.4). The SDS-extraction procedure was continued until all the non-covalently bound material was released. Twenty microlitres of each of the SDS-extractable fractions were analysed for scFv B7 expression via Western blotting (Section 2.3.3) using the primary antibody specific for human kappa light chains (rabbit anti-human kappa light; Table 2.1). The secondary antibody, horseradish peroxidase conjugated goat anti-rabbit IgG was used for detection (Table 2.2).

After all the SDS-extractable material was removed from the cell walls, they were treated with lyticase (a mollusc β-1,6-glucanase) to confirm the scFv B7 antibody fragment was covalently attached to the β-glucan cell wall component. Firstly, the SDS-treated cell walls were washed copious times with 1mM phenylmethylsulphonyl fluoride (PMSF) to remove the unwanted SDS. Forty milligrams of cell wall (wet weight) was resuspended in 40µl of lyticase solution (1.2M sorbitol, 1M sodium phosphate, 3000 units/ml

lyticase). Protease inhibitors pepstatin, antipain, chymostatin and leupeptin were added to the mixture at a final concentration of 30μg/ml. Also, PMSF and EDTA were added to a final concentration of 1mM. The reaction was then mixed gently and incubated with rotation overnight at room temperature. The supernatant, which contained the glucanase extractable material, was collected by centrifugation at 13000 r.p.m. for 5 minutes. Twenty microlitres of the supernatant was resolved by SDS-PAGE (Section 2.3.1) and analysed by Western blotting (Section 2.3.3). The primary antibody specific for the human kappa light chain (Rabbit anti-human kappa light; Table 2.1) followed by the complementary horseradish peroxidase conjugated goat anti-rabbit IgG secondary antibody (Table 2.2). The SDS-treated cell walls from the wild type strain (YMTAB) were incubated with lyticase and analysed by Western blotting for a negative control.

Note – The activity of the lyticase solution was analysed by incubating intact yeast cells in the solution and checking for cell lysis using phase contrast microscopy.

4.3.8 Verification of surface expression of the scFv B7 antibody fragment using an enzyme immunoassay

The *S. cerevisiae* strain, YMTAB carrying the anti-FITC scFv expression vector, pYES2-*SAG1*-scFv B7 was induced with galactose for 24 hours (Section 4.3.3). The yeast cells were analysed for surface expression of the scFv using an enzyme immunoassay (Section 2.3.5). The antibodies used for detection were the rabbit anti-human kappa chain specific primary antibody (Table 2.1) followed by a horseradish peroxidase conjugated goat anti-rabbit IgG secondary antibody (Table 2.2). The wild type strain (YMTAB) transformed with the expression vector, pYES2 was used throughout as a negative control.

4.3.9 Treatment with Factor Xa

After galactose induction, 2 x 10^7 cells expressing the scFv B7 antibody fragment of their surfaces were washed three times in PBS. The cells were collected by centrifugation at 13000 r.p.m. for 2 minutes and resuspended in

50µl of Factor Xa buffer, pH 8.0 (20mM Tris-HCl, 100mM NaCl, 2mM $CaCl_2$). One microgram of the Factor Xa protease and calcium chloride to a final concentration of 2.5mM (Wrighton *et al.*, 1996) were added to the cell suspension. Additionally, a second reaction without any Factor Xa protease was set up and ran concurrently as a control. All reactions were incubated with rotation for 24 hours at room temperature and analysed for scFv expression using an enzyme immunoassay (Section 2.3.5). The antibodies using for detection were identical to Section 4.3.8 and cycloheximide at a concentration of 10µg/ml was present in all solutions throughout the treatment and analysis of the Factor Xa experiment.

4.3.10 Indirect immunofluorescence

The YMTAB strain harbouring the anti-FITC single chain Fv expression vector, pYES2-*SAG1*-scFv B7 was induced for 24 hours with galactose (Section 4.3.3). The enzyme immunoassay protocol (Section 2.3.5) was then carried out on these cells, except instead of detecting specific binding using the chromogenic substrate, 3, 3', 5, 5'-tetramethylbenzidine (TMB), the cells were viewed under an epi-fluorescence microscope. Firstly, the cells were incubated with the primary antibody specific for human kappa light chains (rabbit anti-human kappa; Table 2.1) followed by the FITC-conjugated goat anti-rabbit IgG secondary antibody (Table 2.2). After treatment with the secondary antibody, the cell pellet was resuspended in 50 µl of mounting medium that contained the anti-bleaching agent, p-phenylenediamine. The cells were then viewed under a Leica DMRB epi-fluorescence microscope using the appropriate filter block (490-495nm). Photographs were taken at a magnification of x400 using a Leica MP548 automatic exposure meter and Kodak Ektachrome 400 film.

Note: A YMTAB strain carrying the control plasmid, pYES2 was prepared as a negative control.

4.3.11 Functional identification of the anti-FITC antibody fragments on the surface of *S. cerevisiae*

After galactose induction (Section 4.3.3), approximately 2 x 10^7 cells expressing antibody fragments (scfv or Fab) on their cell surfaces were washed three times with PBS. Fifty microlitres of 0.2μM filtered FITC-BSA at a concentration of 1.5 x 10^{-6}M was added to the cell pellet. The pellet was then gently resuspended in the antigen and incubated for 1 hour at room temperature on a vertically rotating platform. All samples were kept in the dark (wrapped in aluminium foil) for this incubation time. The cells were then collected by centrifugation at 13000 r.p.m. for 30 seconds and the cell pellet was washed five times with PBS. The cells were then resuspended in a 100μl of mounting medium containing the anti-bleaching agent p-phenylenediamine. The cells were then viewed under an Olympus BH2 epi-fluorescence microscope using the appropriate filter block (490-495nm). Photographs were taken using at a magnification of x400 using an automatic exposure meter and Kodak Ektachrome 400 film.

Note: A YMTAB strain carrying the control plasmid, pYES2 was prepared as a negative control.

4.4 RESULTS

4.4.1 Functional expression of anti-FITC Fab fragments A2 and B7 in *E. coli*

E. coli cells containing either vector pUC119His6mycXba-FITC A2 or pUC119His6mycXba-FITC B7 were shown to be expressing functional Fab fragments into their culture supernatant after an induction time of 24 hours (Table 4.1). The supernatant of a wild type *E. coli* culture displayed no fluorescein-5-isothiocyanate binding activity.

The three sets of antibodies used to detect antibody fragments on the surface of *S. cerevisiae* were analysed to confirm their specificity. As expected, the polyclonal antibody specific for human IgG Fab antibody fragments bound precisely to the anti-FITC Fab fragments, B7 and A2. The light chain of the anti-FITC Fab A2 antibody fragment comprised of the

human variable (DPL-3) and constant lambda domains. Therefore, the monoclonal antibody raised against the human lambda light chain was shown to bind specifically to the anti-FITC Fab A2 antibody fragment only. Additionally, the polyclonal antibody specific for the human kappa light chains bound exclusively to the anti-FITC Fab B7 antibody fragments due to the fact its light chain consisted of the human variable (DPK-12) and constant kappa segments.

Table 4.1 - ELISA results obtained from the supernatants of IPTG-induced cultures of TG1[pUC119HismycXba-FITC A2] and TG1[pUC119HismycXba-FITC B7].

Fab clone	Antibodies specific for human: IgG1 Fab fragments	Lambda light chain	Kappa light chain
	Abs_{450}	Abs_{450}	Abs_{450}
FITC-A2	1.71,1.62 (1.67)	1.73,1.61(1.67)	0.03, 0.03 (0.03)
FITC-B7	1.72,1.45 (1.59)	0.01, 0.00 (0.01)	1.64, 1.71(1.68)
TGI cells only	0.34,0.40 (0.37)	0.09, 0.04 (0.07)	0.05, 0.04 (0.05)

The values in brackets represent the mean of the two individual determinations.

4.4.2 Optimisation of galactose induction time

A control experiment monitoring β-galactosidase activity in a culture after induction with galactose was performed to deduce the optimum galactose incubation time.

The supplied control expression vector, pYX243-LacZ, was transformed into the *S. cerevisiae* strain, YMTAB and leu^+ transformants were PCR screened to confirm the plasmid was present. A single positive transformant was grown to mid-exponential phase (5 x 10^6 cells/ml) in defined medium containing the non-catabolite repressing carbon source, raffinose. Galactose was added to a recommended final concentration of 2% (Adams *et al.*, 1996) and samples from the culture were analysed for β-galactosidase activity at regular intervals.

Expression of β-galactosidase was almost immediate, due to the fact that the culture was grown in a non-catabolite repressing carbon source. Cells grown in glucose before exposure to galactose exhibit a long lag phase of 3-5 hours before expression is induced (Adams, 1972). Monitoring the β-galactosidase activity over galactose incubation time illustrated that expression gradually increased to a maximum approaching twenty-four hours. After this time, β-galactosidase activity slowly decreased and reached a plateau after approximately thirty-two hours.

4.4.3 Detection of the anti-FITC Fab A2 antibody fragments on the surface of *S. cerevisiae*

A) Western analysis of cell wall extracts

Cells potentially expressing the anti-FITC Fab A2 antibody fragment on their surfaces were broken with glass beads. Cell walls were isolated and were found to represent approximately 40% of the total cell weight (wet weight). Cells walls have been reported to account for roughly 30% of the dry weight of a cell (Lu *et al.*, 1995).

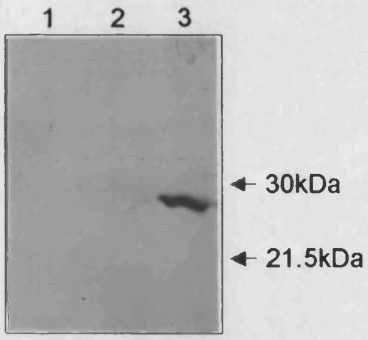
The light chain of the Fab A2 antibody fragment was located in the SDS-extractable material of the cell walls by Western analysis. Both the monoclonal antibody specific for the human lambda chain (Figure 4.1a) and the serum raised to human IgG1 Fab fragments (Figure 4.1b) detected the light chain at the expected molecular weight of approximately 28kDa.

Numerous non-specific bands were observed on the Western blot using the serum raised against anti-human Fab antibody. It was presumed that antibodies in the serum were binding to epitopes on native yeast SDS-extractable cell wall proteins and producing these non-specific results. Some of the antibodies in this serum were expected to bind to the heavy chain-agglutinin protein fusion of the Fab A2 molecule. The molecular weight of the heavy chain-agglutinin fusion was envisaged to be above 70kDa due to N- and O-glycosylation of the α-agglutinin anchorage domain (Section 4.1.1). A high molecular weight smear at approximately 200kDa was observed which could represent the highly heterogenous glycosylated form of the heavy chain-agglutinin fusion. To confirm this result, further analysis with an

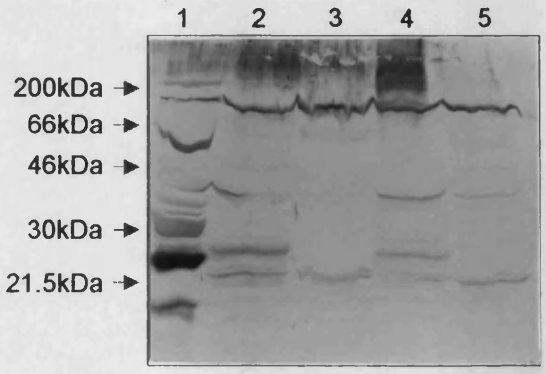
antibody specific for the human heavy chain was performed, but no appropriate bands at the correct molecular weight were observed.

Figure 4.1 - Western analysis of the SDS-extractable cell wall fractions obtained from cells expressing the anti-FITC Fab A2 antibody fragment. The applied fractions represent approximately 12mg of cell walls (wet weight). The periplasmic lysate from *E. coli* strain expressing the anti-FITC Fab A2 was used as a positive control.

a) Western analysis was performed using a primary monoclonal antibody specific for the human lambda light chain. Lane 1, first SDS fraction from wild type strain, YMTAB; Lane 2, first SDS fraction from isolated cell walls expressing Fab A2; Lane 3, positive control.



b) The Western blot was incubated with an anti-human Fab IgG antiserum. Lane 1, positive control; Lane 2, first SDS fraction from isolated cell walls expressing Fab A2; Lane 3, first SDS fraction from the wild type strain control (YMTAB); Lane 4, second SDS fraction from isolated cell walls expressing Fab A2; Lane 5, second SDS fraction from the wild type strain control (YMTAB).



B) Individual identification of the Fab light and heavy chains on the surface of S. cerevisiae

Enzyme immunoassays were performed on 2 x 10^7 galactose-induced cells to determine if the individual light and heavy chains of the anti-FITC Fab antibody fragment were present on the surface of *S. cerevisiae* (Table 4.2). An average absorbance of 0.27 and 0.15 above background were observed for the separate light and heavy Fab chains, respectively. These results confirm that both the heavy and light chains were on the surface of *S. cerevisiae*, but the expression levels compared to the anti-FITC scFv B7 (Table 4.3) were considerably lower.

Table 4.2 - Enzyme immunoassay results verifying the expression of the individual light and heavy Fab chains on the surface of *S. cerevisiae*.

Strain	Antibodies specific for human:	
	Lambda light chains	Heavy chains
	Abs_{450}	Abs_{450}
YMTAB[pYX-Light A2 + pYES2-*SAG1*-Heavy A2]	0.37, 0.35, 0.38 (0.37)	0.43, 0.45, 0.47 (0.45)
YMTAB	0.12, 0.12, 0.06 (0.10)	0.31, 0.29, 0.29 (0.30)

The values in brackets represent the mean of the three independent determinations.

C) Functional detection of anti-FITC Fab fragments using fluorescence microscopy

Yeast cells expressing the Fab light and heavy chains were incubated with the fluorescent antigen, fluorescein-5-isothiocyanate. After extensive washing, the cells were viewed using epi-fluorescence microscopy. Theoretically, cells expressing the functional anti-FITC Fab fragment on their cell surfaces should bind the fluorescent antigen. Unfortunately, the degree of fluorescent labelling observed on the surface of the cells expressing the Fab light and heavy chains was very similar to the wild-type control. This result indicated that the individually expressed light and heavy chains were

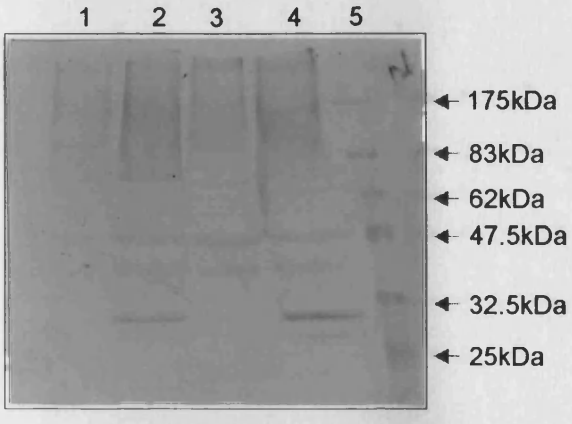
unable to form functional anti-FITC Fab fragments on the surface of *S. cerevisiae*.

4.4.4 Identification of the anti-FITC single chain Fv B7 on the surface of *S. cerevisiae*

A) Covalent attachment to the cell wall

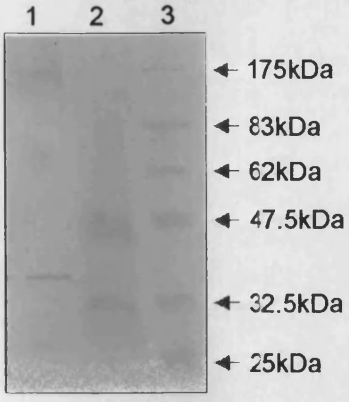
Cells potentially expressing the scFv-agglutinin protein fusion on their cell surfaces were disrupted with glass beads and separated into intracellular and cell wall fractions. The cell walls were treated with hot SDS and the SDS-extractable material was analysed by Western blotting (Figure 4.2). A specific band at approximately 30kDa was detected in the first five SDS extracts. It was presumed that this band represented the scFv B7 part of the protein fusion and the α-agglutinin anchorage domain had been removed by proteolysis. Additionally, a high molecular weight smear (70kDa to >200kDa) was observed to be more reactive in the samples expressing the scFv-agglutinin fusion compared to the wild type control. It was postulated that this smear could represent a highly heterogenous glycosylated form of the scFv-agglutinin protein fusion.

Figure 4.2 - Western analysis of the SDS-extractable cell wall fractions obtained from cells expressing the anti-FITC scFv B7 antibody fragment. The blot was incubated with an anti-human kappa chain antiserum. The applied fractions represent approximately 20mg of cell walls (wet weight). Lane 1, second SDS fraction from the wild type strain, YMTAB; Lane 2, second SDS fraction from isolated cell walls expressing scFv B7; Lane 3, first SDS fraction from the wild type strain, YMTAB; Lane 4, first SDS fraction from the isolated cell walls expressing the scFv B7; Lane 5, Molecular weight protein markers (Broad range).



The SDS-extracted cell walls were treated with glucanase to investigate whether or not the scFv-agglutinin fusion was covalently attached to the cell wall's β-glucan. After numerous unsuccessful attempts, a large diffuse smear was observed between approximately 28 to 83kDa (Figure 4.3). In this smear, there were prominent bands at approximately 30 and 40 kDa, which could represent the individual scFv and α-agglutinin components of the protein fusion, respectively. As mentioned above, these fragments could have been generated by proteolysis, which may have occurred in the long enzymatic incubation step. The expected high molecular weight band (>200kDa) for the glycosylated form of the scFv-agglutinin fusion was not observed. Hyperglycosylation and/or proteolysis may have affected the detection of the entire scFv-agglutinin protein fusion either by the loss of available epitopes or by extensive protein degradation, respectively.

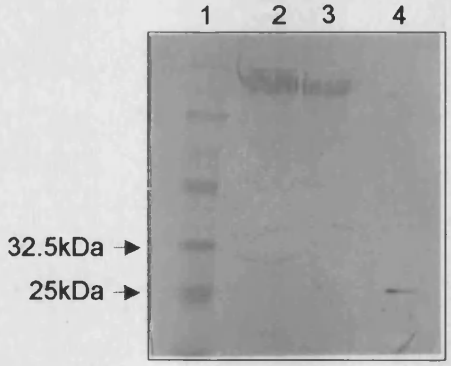
Figure 4.3 - Western analysis of the glucanase-extractable material from SDS-treated cell walls expressing the anti-FITC scFv B7. The blot was incubated with a primary antibody specific for human kappa chains and each lane represents approximately 30mg of SDS-treated cell walls. Lane 1, glucanase-extracted material from the wild type SDS-treated cell walls (YMTAB); Lane 2, glucanase-extracted material from SDS-treated cell walls expressing the scFv B7; Lane 3, Molecular weight protein markers (Broad range).



B) Detection of scFv B7 in the culture supernatant

Culture supernatants were obtained from cells expressing the anti-FITC scFv and a wild type strain harbouring the control vector, pYES2. Both supernatants were concentrated five-hundred fold using ultrafiltration and analysed by Western blotting (Figure 4.4). A band of approximately 30kDa was observed in the concentrated culture supernatant of 1 x 10^9 cells expressing the scFv antibody fragment. This band was presumed to represent the scFv part of the scFv-agglutinin protein fusion as it was absent in the wild type control. It was speculated that protein degradation of the antibody-agglutinin fusion after cell death might account for the scFv antibody fragment in the culture supernatant.

Figure 4.4 - Western analysis of the concentrated culture supernatants. Each lane represents the culture supernatant from approximately 1 x 10^9 cells and the blot was incubated with a primary antibody specific for human kappa light chains. The periplasmic lysate from *E. coli* cells expressing the anti-FITC Fab B7 was used as a positive control. Lane 1, Molecular weight protein marker (Broad range); Lane 2, concentrated supernatant from cells expressing the scFv B7; Lane 3, concentrated supernatant from cells harbouring the control vector, pYES2; Lane 4, positive control.



C) Identification of the scFv B7 antibody fragment on the surface of S. cerevisiae

An enzyme immunoassay using a primary antibody specific for human kappa light chains was performed on galactose-induced intact cells harbouring the anti-FITC scFv expression vector, pYES2-*SAG1*-scFv B7 (Table 4.3). A mean absorbance (at 450nm) of 0.94 for the three samples expressing the

scFv compared to 0.09 for the three control samples was obtained. This result confirms that scFv B7 antibody fragments are on the surface of *S. cerevisiae* and are accessible to macromolecules.

Table 4.3 - Enzyme immunoassay results verifying the expression of the anti-FITC scFv B7 on the surface of *S. cerevisiae*. The primary antibody specific for human kappa light chains was used for detection.

Strain	Abs_{450}
YMTAB[pYES2-*SAG1*-scFv B7]	0.95, 0.92,0.94 (0.94)
YMTAB	0.09, 0.08, 0.10 (0.09)

Values in brackets represent the mean of the three individual determinations.

D) Treatment with the Factor Xa protease

Galactose-induced cells carrying the anti-FITC single chain Fv surface expression vector, pYES2-*SAG1*-scFv B7, were incubated with the Factor Xa protease for 24 hours. Cells without any Factor Xa were also incubated for 24 hours. After this time, both sets of cells were analysed for surface expression of the scFv B7-agglutinin fusion using an enzyme immunoassay (Table 4.4).

Cycloheximide was used throughout the experiment to stop protein production; therefore the results of the enzyme immunoassay ought to represent the scFv-agglutinin protein fusions which have become accessible directly after galactose induction. It was presumed if the Factor Xa protease was cutting at its particular recognition site (Ile-Glu-Gly-Arg), present in-between the scFv and α-agglutinin anchorage domains in the fusion protein, a decrease in the level of scFv expression would be noted. The absorbance readings for the protease-treated and untreated samples were very similar (Table 4.4). Therefore, these preliminary results indicate that the Factor Xa protease was unable to cut at the cleavage site and the scFv antibody fragment was not released from the cell wall. Since earlier experiments (Section 4.4.4c) indicated that at least some of the scFv antibody fragment was present on the yeast cell surface, it was speculated that the cleavage

site may have been hidden in the outer fibrillar layer of the cell wall and therefore may not have been accessible to the Xa protease.

Table 4.4 - Enzyme immunoassay results investigating the Factor Xa treatment of cells expressing the scFv-agglutinin protein fusion (two experiments).

Strain	Factor Xa	Abs_{450}
YMTAB[pYES2-SAG1-scFv B7]	+	0.68, 0.65
" "	-	0.69, 0.74

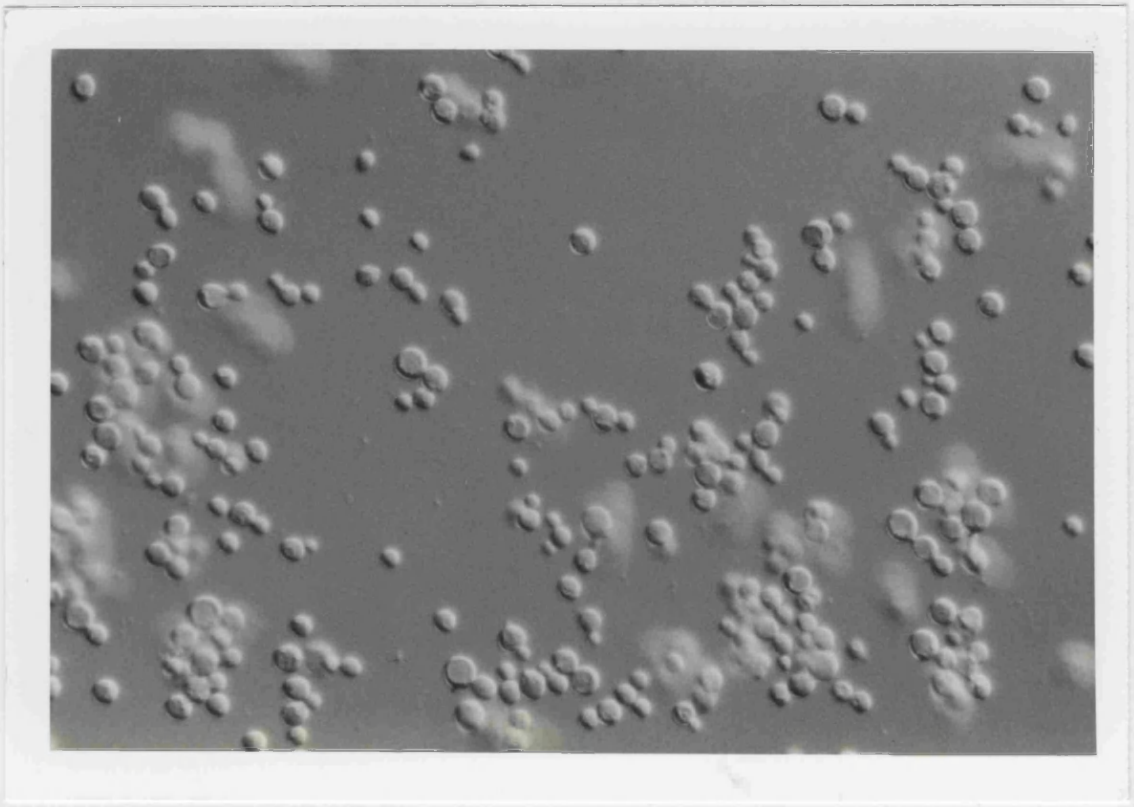
E) Indirect immunofluorescence

The immunofluorescent labelling of cells expressing the scFv-agglutinin protein fusion was achieved using a serum specific for human kappa light chains followed by a complementary FITC-labelled conjugated secondary antibody (Figure 4.5a). All galactose-induced cells harbouring the anti-FITC scFv expression vector were fluorescently labelled but a large variation in fluorescent intensities from cell to cell was observed. The differences in fluorescent intensities were presumed to be due to variation in expression levels caused by the changes in copy number of the surface display vector between cells. Even the very small buds of the cells were labelled which suggests the scFv-agglutinin fusions were continuously incorporated into the yeast cell wall throughout the cell cycle.

No fluorescent labelling was observed on the surface of cells harbouring the control vector, pYES2 (Figure 4.5b). Therefore, it can be assumed the detection antibodies did not react with any native epitopes on the surface of *S. cerevisiae.* These indirect immunofluorescence results support the enzyme immunoassay results (Section 4.4.4c) and imply the scFv B7 antibody fragments are present in the outer fibrillar layer of the cell wall.

Figure 4.5 - Indirect immunofluorescence results using a primary antibody specific for the human kappa chain followed by an FITC-conjugated goat anti-rabbit IgG serum. Nomarsky image (top) and matching fluorescent photograph (bottom). Magnification: x 400.

a) Galactose-induced cells harbouring the scFv B7 expression vector, pYES2-*SAG1*-scFv B7.



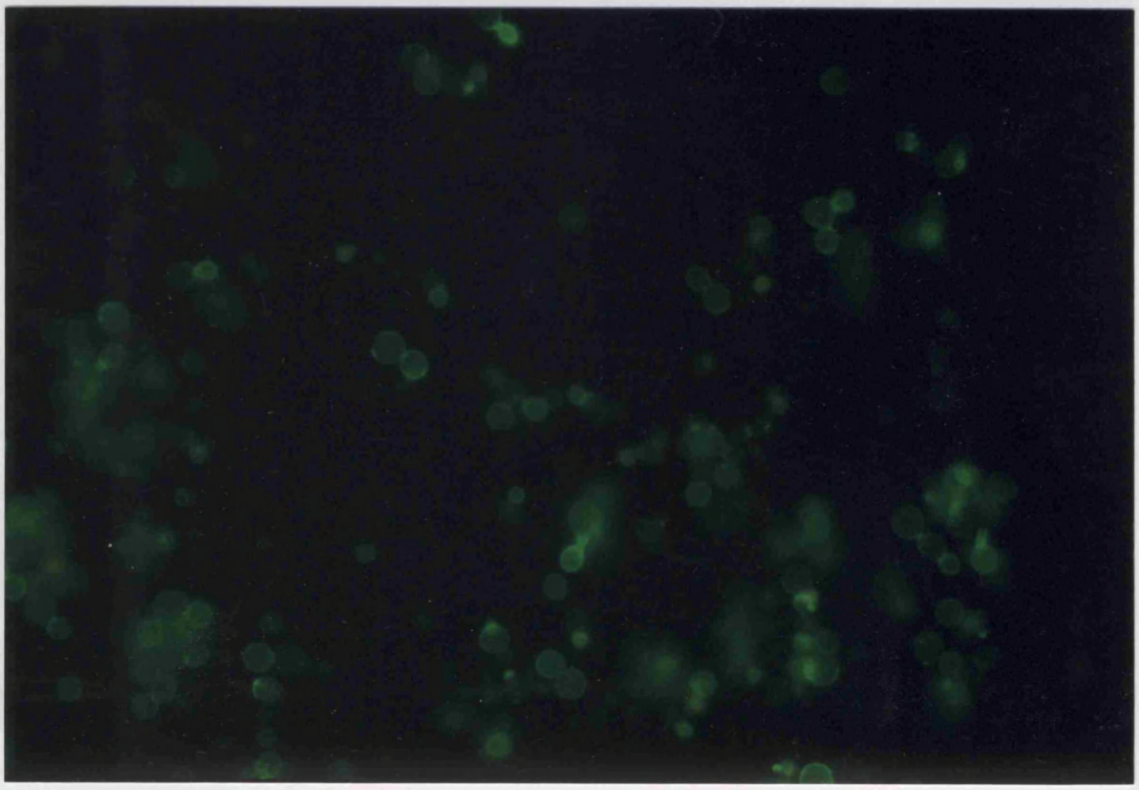
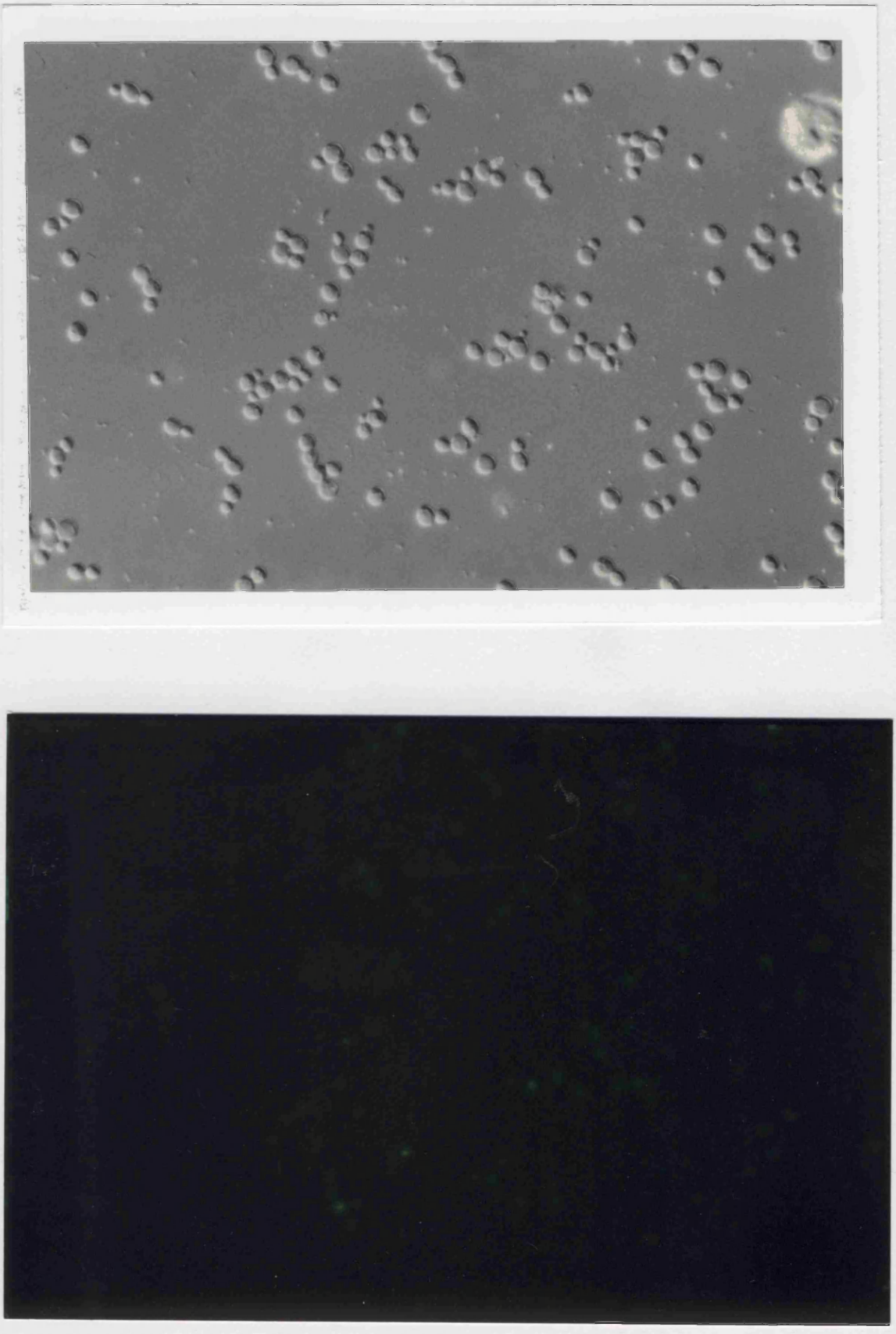


Figure 4.5 cont:

b) Galactose-induced cells carrying the control expression plasmid, pYES2.



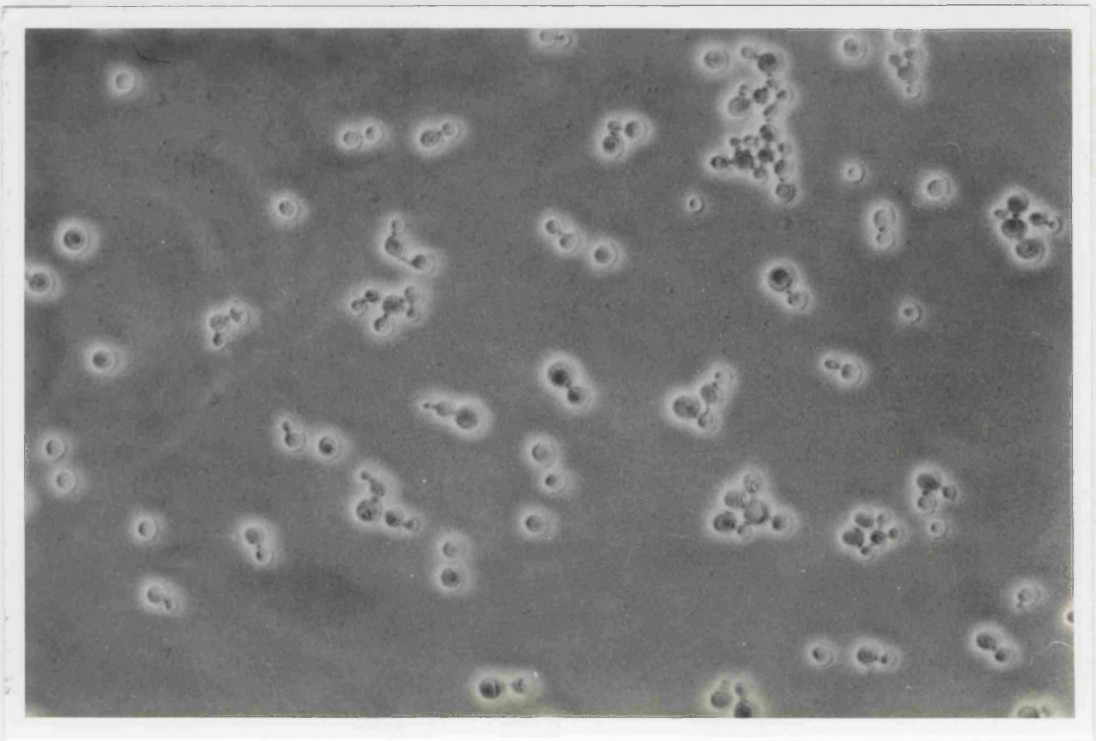
F) Functional identification of the scFv B7 using the antigen FITC-BSA

Fluorescence microscopy was used to illustrate that functional anti-FITC single chain Fv antibody fragments were on the exterior surface of *S. cerevisiae.* All cells expressing the scFv B7, which had been incubated with the fluorescein-5-isothiocyanate antigen, were fluorescently labelled (Figure 4.6a). As expected, the level of fluorescence varied greatly from cell to cell due to the differences in expression levels. The wild type cells harbouring the control plasmid, pYES2, were barely detectable when observed by fluorescence microscopy (Figure 4.6b).

These results indicate the single chain Fv's antigen-binding site is functional and accessible to large macromolecules on the surface *of S. cerevisiae.*

Figure 4.6 - Fluorescence microscopy (bottom) and phase contrast images (top) of yeast cells displaying the anti-FITC scFv B7. Cells were incubated with the fluorescein-5-isothiocyanate antigen. Magnification: x 400.

a) Galactose-induced cells harbouring the scFv B7 expression vector, pYES2-*SAG1*-scFv B7.



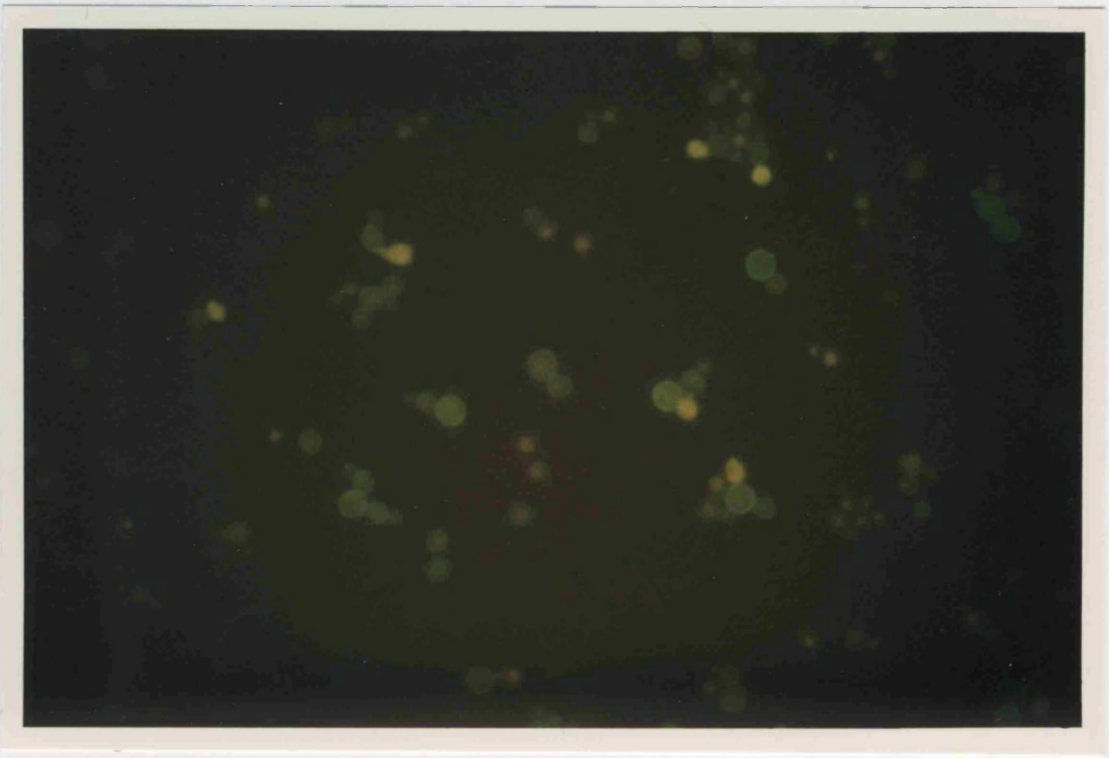
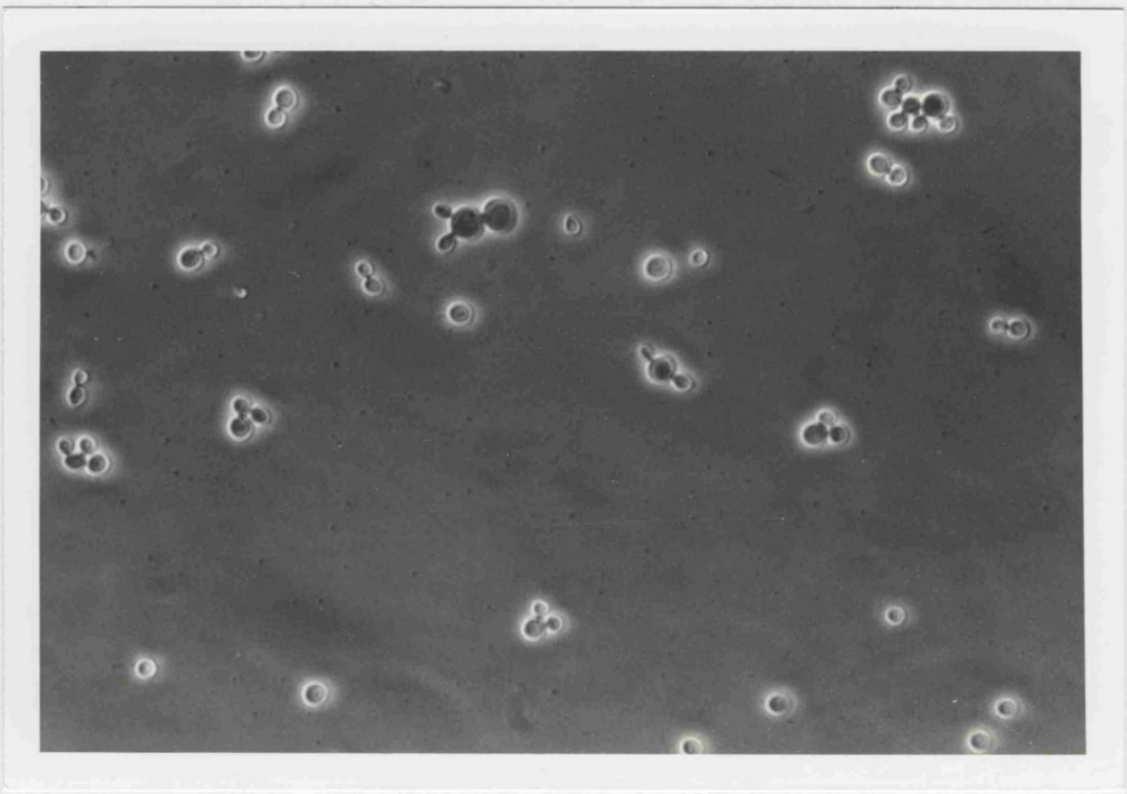
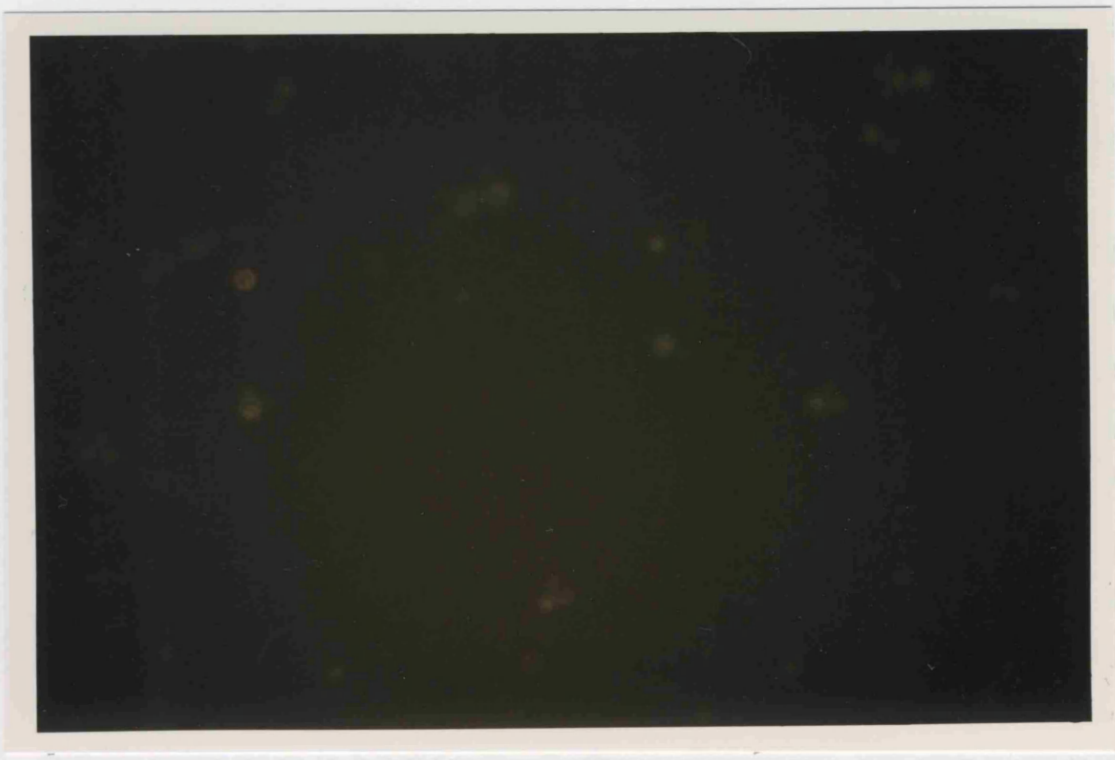


Figure 4.6 cont:

b) Galactose-induced cells carrying the control expression plasmid, pYES2.





4.5 DISCUSSION

The anti-FITC Fab A2 and B7 antibody fragments (Griffiths *et al.*, 1994) were shown to be functionally expressed in the bacterial supernatant (Table 4.1). These ELISA results and the analysis of the DNA sequences (Section 3.4) determined that the *E. coli* expression vectors, pUC119HismycXba-FITC A2 and pUC119HismycXba-FITC B7, were correctly supplied.

The optimum galactose induction time for maximum expression using the *S. cerevisiae* strain, YMTAB was unknown. Therefore, a control expression vector containing the β-galactosidase reporter gene was transformed into YMTAB and its activity was monitored over time after galactose induction. An optimum galactose incubation time of approximately 24 hours was established and all further expression experiments were carried out over a 24-hour time period for consistency.

4.5.1 Detection of antibody fragments on the surface of *S. cerevisiae*

Antibodies cannot penetrate the yeast cell wall (Pringle *et al.*, 1991), therefore the enzyme immunoassay results obtained for both the light and heavy Fab chains (Section 4.4.3b) and the single chain Fv (Section 4.4.4c) proved that these antibody fragments were on the surface of *S. cerevisiae.* In the case of the single chain Fv, additional results acquired from the direct (Section 4.4.4f) and indirect immunofluorescence (Section 4.4.4e) experiments reinforced this conclusion.

Cells expressing the anti-FITC scFv B7 were shown to be specifically binding the antigen, fluorescein-5-isothiocyanate, on their cell surfaces (Section 4.4.4f). Hence, it can be assumed that the single chain Fv's antigen-combining site was functional on the yeast's surface. It was unknown how many scFv molecules on the cell surface were forming functional antigen-combining sites. An additional quantitative double labelling experiment using a second fluorophore, which is excited at a different wavelength to fluorescein-5-isothiocyanate and can recognise an unmutated part of the scFv antibody fragment, could have been used to compare the level of surface expression to functionality. Furthermore, examination of the single or

doubly labelled yeast cells by laser-scanning confocal microscopy would have given additional information about the surface distribution of the functional antibody fragments.

The indirect and direct immunofluoresence results obtained for cells expressing the single chain Fv antibody fragment determined that the level of surface expression varied widely from cell to cell (Sections 4.4.4e and 4.4.4f). It was presumed that the differences in the level of expression was due to the variance in the copy number of the expression vector and this problem could be overcome by integrating the antibody DNA into the α-agglutinin locus (Schreuder *et al*, 1996). Homogenous surface expression is essential for the development of an antibody selection procedure. It was intended that a successful selection process that can discriminate between cells that are expressing high and low affinity antibodies on their surfaces would be developed in the next part of the project (Part II).

Unfortunately, anti-FITC Fab antibody fragments were not functionally identified on the surface of *S. cerevisiae* (Section 4.4.3c). The reasons for this unsuccessful outcome could be numerous. The experimental strategy of this part of the project relies on disulphide bond formation between the Fab heavy and light chains in the endoplasmic reticulum for a functional anti-FITC Fab antibody fragment to be created. The light chain expression construct, pYX-Light A2, contains an α-agglutinin signal sequence that targets expression to the secretory pathway and the attachment of the light chain to the cell wall is reliant on the heavy chain. The heavy chain surface display construct, pYES2-*SAG1*-Heavy A2, includes the molecular information encoded by the α-agglutinin gene, therefore should covalently attach the formed Fab antibody fragment to the glucan of the cell wall. If the Fab molecule does not form, the light chain should be secreted into the culture supernatant (Wojciechowicz *et al.*, 1993). Enzyme immunoassay results (Section 4.4.3b) indicated that the Fab light chain was present at a low level on the yeast's cell surface. Therefore, it was presumed that the light and heavy chains were imprecisely associating and the antigen-combining sites were forming incorrectly. Another factor affecting the detection of anti-FITC Fab fragments on the cell surface could be the permeability of the cell wall

(Klis, 1994). The scFv (~32 kDa) is half the molecular weight of the Fab (~ 57kDa) antibody fragment, therefore the Fab molecule may have had more difficulties passing through the cell wall owing to its size. Further analysis of the direct immunofluorescence results with imaging enhancing equipment may have detected a low level of functional anti-FITC Fab fragments.

4.5.2 Covalent attachment of antibody fragments to the cell wall

An important objective of this chapter was to determine if the antibody fragments were covalently attached to the cell wall's β-glucan. The native α-agglutinin is exclusively extractable with glucanase but fusion proteins created with the C-terminal half of the α-agglutinin gene have been shown to be partly extractable with hot SDS [non-covalently bound (Schreuder *et al.*, 1993)]. Schreuder and colleagues (1993) expressed an α-galactosidase/α-agglutinin anchorage domain fusion protein on the surface of *S. cerevisiae.* In these experiments, 31% of the fusion protein was released with hot SDS and 69% by treatment with glucanase. It was reasoned that the higher level of expression of the fusion protein compared to the native α-agglutinin was the cause of this difference. The SDS-extraction experiments carried out in this Chapter (Section 4.4.3a and 4.4.4a) indicated that a large fraction of the antibody-agglutinin fusion was present in the non-covalently attached fractions, although no quantitative experiments were carried out. The total removal of the non-covalently associated proteins from isolated cell walls expressing the scFv B7 fragment required five concurrent SDS treatments. Previously, it has been reported that two hot SDS treatments are sufficient to remove all the non-covalently bound protein (Schreuder *et al.*, 1993). It was anticipated that a small amount of the antibody-agglutinin construct would be identified in the SDS fractions due to biosynthetic intermediates that are not associated with the glucan such as the GPI anchored plasma membrane and the soluble periplasmic forms (Lu *et al.*, 1995). But the reasons why additional SDS treatments were required are unknown.

To demonstrate the scFv-agglutinin protein fusion was covalently attached to the β-glucan, the SDS-treated cell walls were incubated with an impure β-1-6-glucanase, lyticase (Section 4.4.4a). However, the results of

these experiments were not totally conclusive. After treatment with glucanase, the expected high molecular weight band (>200kDa) representing the highly heterogenous glycosylated form of the scFv-agglutinin was not observed. Nevertheless after several attempts, a lower molecular weight smear of 28-83kDa was identified. Within this smear, two prominent bands of approximately 30 and 40kDa were detected, which theoretically could represent the scFv and α-agglutinin components of the protein fusion. This suggests proteolysis may be affecting these results, although an *S. cerevisiae* strain deficient in two main vacuolar proteases and numerous protease inhibitors were used.

Hyperglycosylation of the fusion proteins may have resulted in the loss of available epitopes, which could have affected detection sensitivity (Romanos *et al.*, 1992). N-linked carbohydrates can be easily removed from a protein using the enzyme, endoglycosidase H (endo H). This enzyme cleaves the N-chains after the first N-acetyl glucosamine residue and removal of the N-chains can be monitored by an increase in electrophoretic mobility due to loss of molecular mass (de Nobel, 1996). Treatment of the glucanase-extractable material with Endo H may have increased the sensitivity of the immunoblotting experiments. Additionally using tunicamycin, a drug that inhibits the addition of the N-chains to a protein or a *mmn9* yeast mutant that produces glycoproteins with truncated N-chains, would have had the same effect. Due to the difficulties experienced with the glucanase treatment of cell walls expressing the scFv B7 antibody fragment, the detection of the Fab heavy chain in the glucanase-extractable material was never attempted.

4.5.3 Conclusions

The aim of this chapter, to express a functional antibody fragment on the surface of *S. cerevisiae,* was achieved. Functional anti-FITC scFv antibody fragments were detected on the surface of the yeast cells by fluorescence microscopy. Unfortunately, the fluorescence-activated cell sorting (FACS) antibody selection system that was devised to separate yeast cells displaying high affinity antibodies from a mixed population of yeast cells was not

established. This was due to the fact that another research group demonstrated the capabilities of a FACS-based antibody selection system first (Boder and Wittrup, 1997). Hence, Part II of this project attempted to develop a novel strategy for the selection of high affinity antibodies on the surface of *S. cerevisiae* using the mating characteristic of agglutinin mutants.

PART II

IMITATING THE GERMINAL CENTRE REACTION BY MATING YEAST

Strategy

During the germinal centre reaction (Section 1.3.3), B-lymphocytes with mutated antigen receptors (centrocytes) are selected on the basis of their capacity to bind antigen which is held in a limited amount on follicular dendritic cells. Binding to follicular dendritic cells is an unconditional prerequisite for centrocyte survival and cells which do not bind die rapidly by apoptosis (reviewed in Maclennan, 1994; Kelsoe, 1996; Camacho *et al.*, 1998). This is the presumed basis of affinity maturation during immune responses and is the selection process this project is aiming to mimic (Figure 1.11).

S. cerevisiae **a**- and α-haploid cells express complementary cell surface receptors known as **a**- and α-agglutinins that mediate cell to cell adhesion during mating (Section 1.6.2). Agglutinins are critical for mating under conditions that do not promote stable cell to cell contact (liquid media) but are not essential under conditions that promote such contact [solid media (Lipke *et al.*, 1989; Roy *et al.*, 1991)]. This characteristic of agglutination is paramount in this proposed strategy, which is as follows:

Firstly, the centrocytes of the germinal centre reaction will be represented by α-cells with a single chain Fv antibody fragment replacing the N-terminal binding domain of the α-agglutinin gene (*SAG1*) to produce an in-frame fusion with the C-terminal α-agglutinin anchorage domain. Thus in the presence of **a**-factor, it is intended that the engineered α-cells will express the antibody fragment on their cell surfaces.

The follicular dendritic cells will be imitated by **a**-cells using the **a**-agglutinin molecule. An antigen gene is intended to be fused to the 3' terminus of the DNA segment encoding the **a**-agglutinin binding domain (*AGA2*) to provide the antigen cell surface display system.

In liquid mating conditions, diploid formation using the engineered agglutinin mutants is dependent on the antigen-antibody interaction. The mating efficiency of these mutants can be easily assessed by using haploid cells that complement each other's nutritional requirements. Therefore, diploid formation can be assayed by plating mated cells onto the appropriate selective media. Cells unable to form a stable cell to cell contact (antigen-antibody interaction) will not form diploids and therefore die because their nutritional requirements are not met. Consequently, α-cells which have antibodies with a moderate to high affinity for the chosen antigen on their cell surfaces are selected.

If this selection procedure was shown to be successful, antibody affinities could then be improved in a step-wise manner. The antibody DNA in the genome of α-cells could be mutated by various mutagenic techniques (i.e. error-prone PCR or a yeast/bacterial mutator strain) and advantageous mutations could then be chosen using the diploid selection procedure. An effective diploid selection process will be a very fast and efficient way of screening a mutant antibody surface display library. By growing the selected diploid cells under nutritional starvation, meiosis will occur and haploid α-cells displaying the beneficially mutated antibody could be isolated. If a yeast mutator strain is used, an 'in-cell' mutational system could be developed which would overcome the necessary DNA manipulations required after antigen-selection when using methods such as error-prone PCR. Therefore, single yeast cells, analogous to the B-cells in the germinal centre reaction, could undertake rounds of mutation and selection.

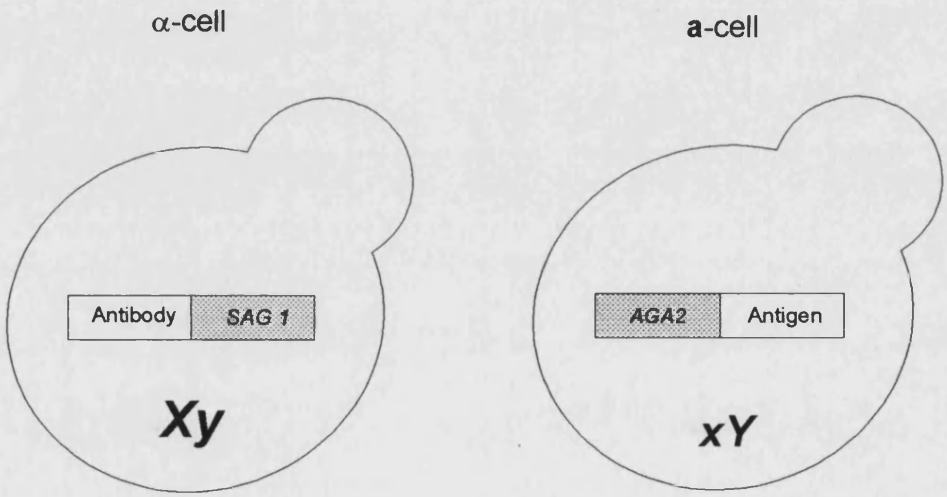
The attempted realisation of this strategy forms the subject of the next two chapters: Chapter 5 deals with the construction of the agglutinin mutants and Chapter 6 details the attempted mating of these mutants to imitate the germinal centre reaction.

Outline of the strategy proposed to imitate the germinal centre reaction:

STEP 1 - *Integration of the antibody and antigen DNA into the appropriate agglutinin loci.*

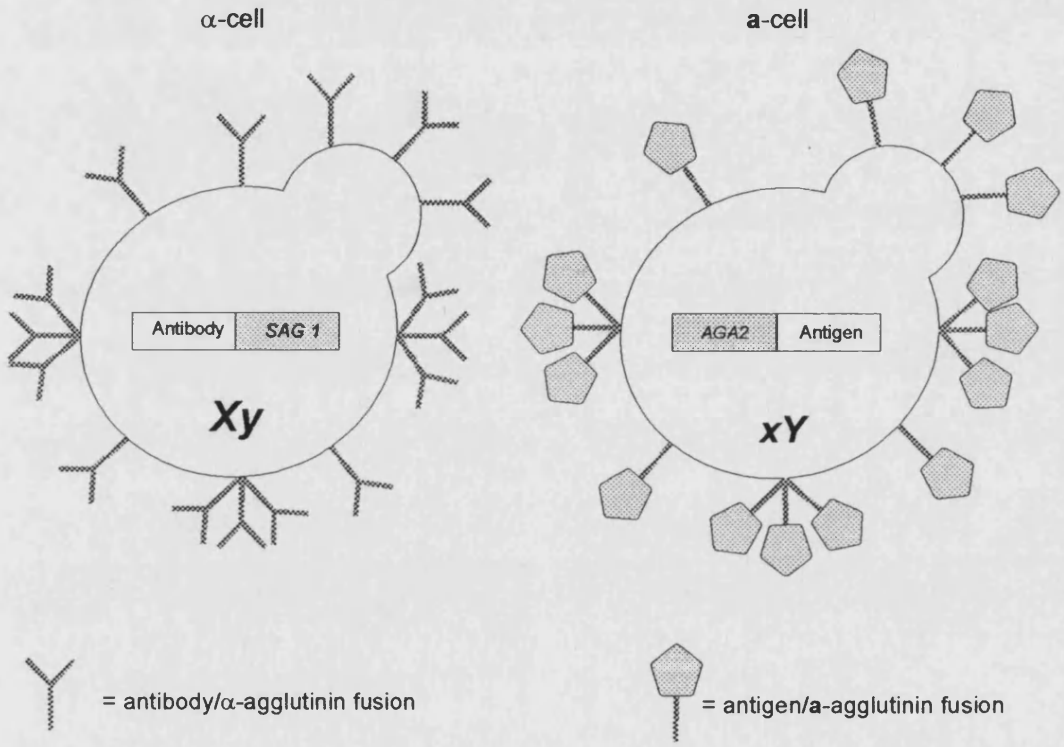
i) Integrate the antibody fragment into the α-cell genome to create an in-frame fusion with the 3' half of the α-agglutinin gene *(SAG1)*. Select an α-strain that has a mutation in gene *Y* and therefore can not synthesise the amino acid, Y.

ii) Integrate the DNA encoding the antigen gene into the **a**-cell genome to generate an in-frame fusion to the 3' terminus of the **a**-agglutinin binding domain gene (*AGA2*). Choose an **a**-strain which has a mutation in gene *X* and therefore can not make the amino acid, X.



STEP 2 - *Mating of the engineered yeast strains*

Mate the engineered yeast strains under conditions that do not promote stable cell to cell contact (liquid media). If the antibody on the surface of the engineered α-cell has a moderate affinity for the antigen displayed on the surface of the **a**-cell they ought to interact to form a diploid.



STEP 3 - *Analysis of diploid formation*

Select for diploids by growing mated cultures on the appropriate selective plates (without amino acids X and Y). Non-mated cells (haploids) die because their nutritional requirements are not fulfilled.

STEP 4 - *Mutation of antibody DNA*

Allow the selected diploid cells to undergo meiosis and isolate the haploid α-cells. Mutate the antibody DNA in the genome of the α-cells and mate with engineered **a**-cells for a further round of antigen selection.

CHAPTER 5

CONSTRUCTION OF AGGLUTININ MUTANTS

5.1 INTRODUCTION

The previous chapter showed that a single chain Fv antibody fragment can be functionally expressed on the surface of *Saccharomyces cerevisiae* using the molecular information of the signal sequence and C-terminal half of the α-agglutinin gene. Immunofluorescence experiments indicated that all cells were expressing the scFv/α-agglutinin protein fusion but the expression level varied widely from cell to cell. Variation of the expression levels was presumed to be dependent on the copy number of the genes present in each cell (Schreuder *et al.*, 1996). Selection of high affinity antibodies from cells with considerably different expression levels would be difficult, because undesirable cells with moderate affinity antibodies but higher expression levels maybe selected. Consequently, an alternative strategy has been designed which overcomes the unstable inheritance of the yeast episomal plasmid by integrating the antibody gene into the genome of *S. cerevisiae*.

Theoretical arguments suggest that the larger and more diverse an antibody library, the higher the probability of finding a high affinity antibody (Perelson and Oster, 1979). In anticipation that a mutant antibody library will eventually be created, the Cre-*lox* recombination system was chosen to integrate the chosen DNA fragments into the appropriate sites in the genome. This approach should overcome the low transformation efficiency achievable by standard integration techniques such as the one-step gene replacement method [1-100 transformants per μg of DNA (Wach, 1996)]. Hypothetically, a Cre recombinase integrated library should only be restricted in size by the transformation efficiency of the exchange vector. Transformation efficiencies in the order of 10^6 transformants per μg of plasmid DNA per 10^8 cells are attainable (Schiestl *et al.*, 1993). Moderate

affinity (μM) antibodies, similar to antibodies from a primary immune response, could possibly be isolated directly if a library of 10^7 colonies was generated (Hoogenboom, 1997). The affinity of the antibody of interest could then be improved in a stepwise fashion by rounds of random mutation and antigen-selection, analogous to the germinal centre reaction (Section 1.3.3).

The background information concerning the Cre-*lox* recombination system and the chosen antibody-antigen interaction is reviewed below:

5.1.1 Cre-*lox* site-specific recombination system

In 1981, N. Sternberg and D. Hamilton discovered the Cre-*lox* site-specific recombination system in bacteriophage P1. It consists of a site, *loxP*, where recombination takes place and a protein, Cre, which carries out the recombination reaction. In the bacteriophage P1 life cycle, the Cre-*lox* system performs at least two functions: it cyclises the viral DNA after its infection into bacteria and it breaks down any DNA dimers that form following replication (Austin *et al.*, 1981). Consequently, the phage genome is maintained as a monomeric, unit-copy plasmid.

The Cre recombinase gene was localised using a series of deletion mutants known to contain the gene. Each of the mutants were screened for Cre activity and as a result a 1.1 kilobase DNA sequence was located which encoded a 343 amino acid (38kDa) protein (Sternberg *et al.*, 1986). Cre is a member of the integrase family of site-specific recombinases and consists of two distinct domains that are separated by a short linker. Both the N- and C-terminal domains bind DNA, but only the C-terminal domain mediates actual strand-exchange between *loxP* sites via a catalytic tyrosine residue (Tyr324). The strand-exchange mechanism occurs in a step-wise fashion through a four-way Holliday junction intermediate, with four recombinases and two *loxP* sites (Guo *et al.*, 1997).

The *loxP* site (34 b.p.) consists of two inverted 13 b.p. repeats separated by an asymmetric 8 b. p. spacer region (Figure 5.1). Cleavage and strand exchange of the recombining *loxP* sites occurs within this spacer region (Hoess and Abremski, 1985). This spacer region dictates whether

excision (*loxP* sites in the same orientation) or inversion (*loxP* sites in inverted orientation) of the intervening DNA sequence occurs after recombination. Intermolecular recombination can also be carried out by Cre, but is less efficient than intramolecular recombination (Abremski *et al.*, 1983).

Hoess and collegues (1986) constructed various mutant *loxP* sites containing single base substitution in the spacer region and discovered that homology was required between recombining *loxP* sites (Table 5.1). By changing the base at position 7 in the *loxP* spacer region, a new *loxP 511* site was created (Figure 5.1).

Figure 5.1 The DNA sequences of the *loxP 511* and *loxP* recognition sites. The single base substitution at position 7 in the *loxP 511* site is indicated in bold.

	REPEAT	SPACER	REPEAT
		1 2 3 4 5 6 7 8	
***loxP* site**	ATAACTTCGTATA	ATGTATGC	TATACGAAGTTAT
***loxP 511* site**		**A**	

Table 5.1 The recombination results of the *loxP* and mutant *loxP 511* sites (Hoess *et al.*, 1986).

loxP sites	*in vivo*[a]	*in vitro*[b]
loxP x *loxP*	100	61
loxP 511 x *loxP*	0	1
loxP 511 x *loxP 511*	100	68

[a] The percentage of colonies that have become sensitive to kanamycin as a result of the excision of the kanamycin resistance gene via Cre-mediated recombination between *loxP* sites.

[b] Percentage of plasmid DNA in which recombination between *loxP* sites has resulted in the formation of two new products.

Due to the simplicity of the Cre-*lox* system i.e. the Cre protein requires no cofactors and the *loxP* site is sufficiently large that it is unlikely to occur

naturally in genomes, this system is being increasingly used to manipulate prokaryotic and eukaryotic genomes. Analysis of the eukaryotic yeast, *S. cerevisiae*, indicated there was no authentic *loxP* sites present in its genome (Sauer, 1992). Applications of the Cre-*lox* system so far include: the engineering of synthetic antibody libraries in phage (Waterhouse *et al.*, 1993; Tsurushita *et al.*, 1996), the removal of selectable markers to generate gene knockouts in yeast and mice (Sauer, 1987; Sauer, 1993), site-directed chromosome translocations in plants and mice (Qin *et al.*, 1994; van Deursen *et al.*, 1995) and recombination-activated gene expression (RAGE) in mice (Lakso *et al.*, 1992).

Cre recombinase in yeast

Sauer (1987) reported Cre recombinase could act upon a yeast chromosome and therefore showed Cre was functional in an eukaryotic nucleus. In these experiments, the Cre recombinase gene was placed under the control of the yeast *GAL1* inducible promoter and expression was regulated by the carbon source. A selectable marker (*LEU2*) flanked by two *loxP* sites was placed on two different chromosomes in different orientations. The ability of the cells to become leu$^-$ was shown to be due to *LEU2* excision via functional Cre recombinase. No nuclear localisation sequence was used, therefore the ability of Cre to enter the eukaryotic nucleus is believed to be a consequence of its size. The nuclear pore has a functional radius of 45-55 Å and hence, in principle, the passage of the spherical proteins as large as 50-60 kDa is possible (Barnes and Rine, 1985). Successful multiple gene disruptions using a *loxP-kanMX-loxP* reusable disruption cassette and Cre recombinase have also been reported (Güldener *et al.*, 1996).

Additionally, a yeast multipurpose transposon-based system containing a *loxP* and a *loxR* site flanking the coding regions of either β-galactosidase or green fluorescent protein has been developed (Ross-Macdonald *et al.*, 1997). The *loxR* site was identified in a study investigating recombination between the *loxP* site on the bacteriophage P1 plasmid and the *loxB* site on the bacterial chromosome (Sternberg *et al.*, 1981). Recombination between

the *loxP* and *loxB* sites results in the integration of the P1 plasmid DNA into the bacterial chromosome and the generation of two new hybrid sites, *loxR* and *loxL*. Analysing the recombination frequency between the *loxR* and *loxP* in bacteriophage λ showed that they can recombine at an efficiency 20- to 100-fold lower than two homologous *loxP* sites. The *loxP* and *loxR* sites were chosen for the yeast transposon system because they only contain 16 identical base pairs and therefore should overcome any problems with spontaneous excision of the transposon. This multipurpose transposon system allows the rapid generation and analysis of yeast disruption mutants. After phenotype analysis, the inserted transposon can be reduced by Cre-mediated site-specific recombination to a smaller element that leaves an epitope tag which can be utilized for immunodetection purposes.

5.1.2 The scFv D1.3-HEWL antibody-antigen interaction

The antigen-antibody interaction chosen for this strategy is that between the single chain antibody fragment D1.3 (scFv D1.3) and hen egg white lysozyme (HEWL). The affinity constant for the scFv D1.3-HEWL interaction [Ka = 3 x 10^8 M^{-1}(Hawkins *et al.,* 1993)] is similar to the **a**-/α-agglutinin interaction [Ka = 1 x 10^9 M^{-1}(Lipke *et al.*, 1987)] and therefore should mimic the agglutinin function.

Single-chain Fv D1.3

The monoclonal anti-HEWL D1.3 antibody was derived by immunizing Balb/c mice with lysozyme three times over a period of 36 days and was obtained as part of a project attempting to define antigen-antibody interactions at the molecular level (Harper *et al*, 1987). Its affinity (Ka = 2.7 x 10^8 M^{-1}) is characteristic of a good antibody-antigen interaction from a secondary immune response (Foote and Winter, 1992). The D1.3 antibody contains five non-silent mutations (VH - Ser^{56}→Asn; VH - Glu^{86}→ His; VL - Asn^{50}→ Tyr; VL - Ala^{51}→ Thr; VL - Lys^{52} → Thr; Figure 5.2a) and is believed to originate from the germline sequences PJ14 (VH) and K2 [VK (Hawkins *et al.*, 1993; England *et al.*,1999)]. The somatic mutations are mainly situated

in the complementarity determining regions and increase the affinity of the antibody by 60-fold. Most of the affinity improvement is due to a single residue change (VL - Asn^{50} → Tyr) which is found at the D1.3-lysozyme interface (England *et al.*, 1999).

The scFv of antibody D1.3 was constructed by joining the VH and VL domains together using a flexible linker, $(Gly_4Ser)_3$ and was the first scFv antibody fragment to be functionally expressed on the surface of phage (McCafferty *et al.*,1990). Its affinity constant of 3 x 10^8 M^{-1} (Hawkins *et al.*, 1993) is similar to the affinity of the parent antibody [Ka = 2.7 x 10^8 M^{-1} (Foote and Winter, 1992)] which indicates the scFv D1.3 antibody fragment binds tightly to the antigen, lysozyme.

The antibody fragments of D1.3 have been used in numerous studies including being 'humanised' for therapy (Verhoeyen *et al.*, 1988; Foote and Winter, 1992) and random mutated for a phage display affinity maturation strategy (Hawkins *et al.*, 1993). Consequently, scFv D1.3 is one of the most well characterised antibody fragments available.

Hen egg white lysozyme

Lysozymes are bacteriolytic enzymes that occur widely in cells and secretions of vertebrates. Hen egg white lysozyme is a small, well-characterised enzyme consisting of a single polypeptide chain of 129 amino acids (14.6 kDa). It destroys bacterial cell walls by hydrolysing the β(1→4) glycosidic linkages in the alternating N-acetyl-muramic acid/N-acetylglucosamine polysaccharide component of the cell wall peptidoglycans. Hen egg white lysozyme was the first enzyme to have its three-dimensional structure determined (Blake *et al.*, 1965) and is therefore very well understood structurally.

The HEWL DNA sequence was obtained from a cDNA isolate (Jung *et al.*, 1980) and is encoded by four exons, each of which carries a main structural element. Exon 2 and 3 determine the active site, exon 1 encodes for the translational signal sequence and exon 4 increases the stability of the molecule.

There have been numerous reports of different lysozymes, including HEWL, being expressed in *Saccharomyces cerevisiae* under the control of different yeast promoters and leader sequences (Oberto and Davison, 1985; Jigami *et al.*, 1986; Castañón *et al.*, 1988). All cases claim that the lysozyme is correctly processed, biologically active and has no deleterious effects on the yeast cells. Human lysozyme has also been expressed in *E. coli* cells, but inactive enzyme with incorrect disulphide bonds was found in inclusion bodies (Muraki *et al.*, 1985). Hen egg white lysozyme has been studied extensively in order to understand the role of glycosylation (Kato *et al.*, 1994a; Ueda *et al.*, 1996) and protein stability (Kato *et al.*, 1994b) and hence is one of the most well-defined antigens in the literature.

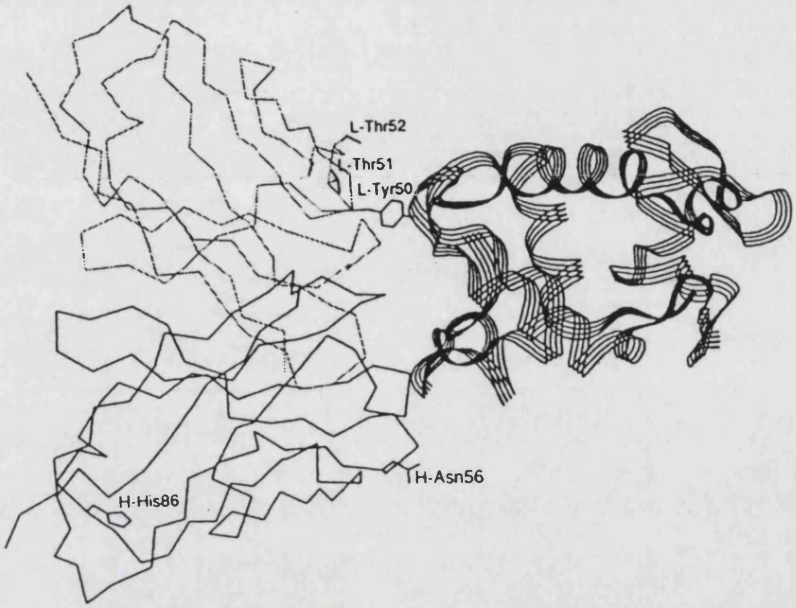
D1.3-HEWL complex

The structural features of the D1.3 antibody (Figure 5.2b) were established initially with the X-ray crystal structure of the Fab fragment complexed with the HEWL at 2.8Å resolution (Amit *et al.*, 1986) and at 2.5Å resolution (Fischmann *et al.*, 1991). Further structural features have also been obtained from high resolution crystal structures of the free D1.3 Fv fragment (Bhat *et al.*, 1990) and the Fv-HEWL complex (Bhat *et al.*, 1994).

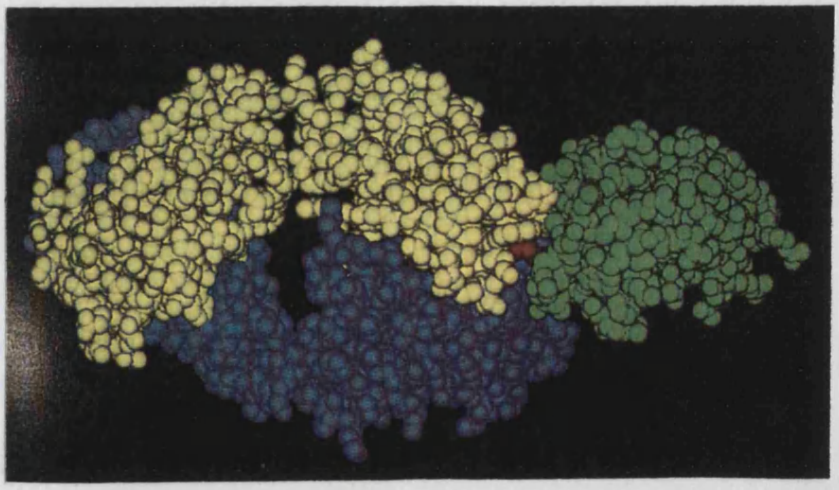
The antibody-antigen interface is tightly packed, with 16 lysozyme and 17 antibody residues making close contact. The interacting surfaces are complementary, with protruding side chains of one lying in the depressions of the other. Many water molecules are observed at the antibody-antigen interface and are believed to play an important role in the stabilization of the complex (Bhat *et al.*, 1994). The lysozyme antigenic determinants recognised by D1.3 consist of two stretches of polypeptide chain: residues 18 to 27 and 116 to 129. All complementarity-determining regions (CDRs) of the D1.3 antibody interact with the lysozyme, especially the variable heavy CDR3. In fact, the VH domain has been expressed individually in *E. coli* and was found to bind to the lysozyme at an affinity only 10-fold weaker (Ward *et al.*, 1989). The D1.3-HEWL complex is one of the most well understood antibody-antigen interactions in the literature and therefore D1.3 antibody fragments are ideal candidates for further mutagenic studies.

Figure 5.2 The D1.3-HEWL complex.

a) The structural location of the five non-silent mutations of antibody D1.3. The α-carbon traces of the variable fragments of Ab D1.3 (VH, solid lines; VL, dashed lines) and of HEWL (ribbon representation) are shown. Reproduced from England *et al.*, 1999.



b) Space-filling representation of the Fab D1.3-HEWL complex. The antibody heavy chain is shown in blue, the light chain in yellow, lysozyme in green and Gln^{121} in red. Reproduced from Amit *et al.*, 1986.



5.1.3 Experimental design for the construction of the agglutinin fusion mutants

Figures 5.3 and 5.4 illustrate the construction of the scFv D1.3/α-agglutinin and the HEWL/**a**-agglutinin gene fusions, respectively. The first stage in the construction of the HEWL and scFv D1.3 agglutinin mutants was to integrate a selectable marker, flanked either side by a *loxP* or *loxP 511*

site, into the agglutinin loci of the respective genomes. Special consideration in the auxotrophic mutations of the haploid yeast strains to be manipulated was taken. The mutations were complementary to the selectable markers of the vectors used in these experiments and also compatible with the diploid selection strategy (Chapter 6).

Classically, null mutants of a particular locus are produced by using a one-step gene transplacement method (Rothstein, 1983). This involves placing small DNA fragments of the gene of interest either side of the selectable marker. After transformation into yeast, homologous recombination between the flanking regions results in a deletion of the gene and the simultaneous integration of the marker gene. Yeast contains very efficient mechanisms for homologous recombination and it is often possible to reduce the flanking DNA to 35-45 b.p. This sometimes leads to problems with mis-integration due to occurrence of sequence polymorphisms in strains with different genetic backgrounds.

In these experiments, the agglutinin disruption cassettes with several hundred base pairs of flanking DNA were constructed to overcome the mis-integration problem and also increase the transformation efficiency (Wach, 1996). Long-flanking homology PCR (LFH-PCR) disruption cassettes were made by linking two PCR-amplified fragments (~1000 b.p. each) produced from genomic DNA of the target locus to the marker gene via a 'splicing by overlap extension' reaction (Horton *et al.*, 1989). Oligonucleotides used to amplify these PCR-amplified fragments were also designed to incorporate the *loxP* and *loxP 511* recognition sites. Hence after transformation, the *loxP/loxP 511* sites were either side of the selectable marker at the relevant agglutinin locus.

The selectable marker chosen for this project was the heterologous dominant resistance marker, *kanMX4*, which encodes for an aminoglycoside phosphotransferase that permits selection of transformants resistant to the drug geneticin (G418). Wach *et al.* (1994) constructed this module which contains the kan^r opening reading frame of the *E. coli* transposon Tn*903* fused to the transcriptional and translational control sequences of the *TEF* gene of the filamentous fungus *Ashbya gossypi* (Figure 5.6a). This module can be transformed into any strain and

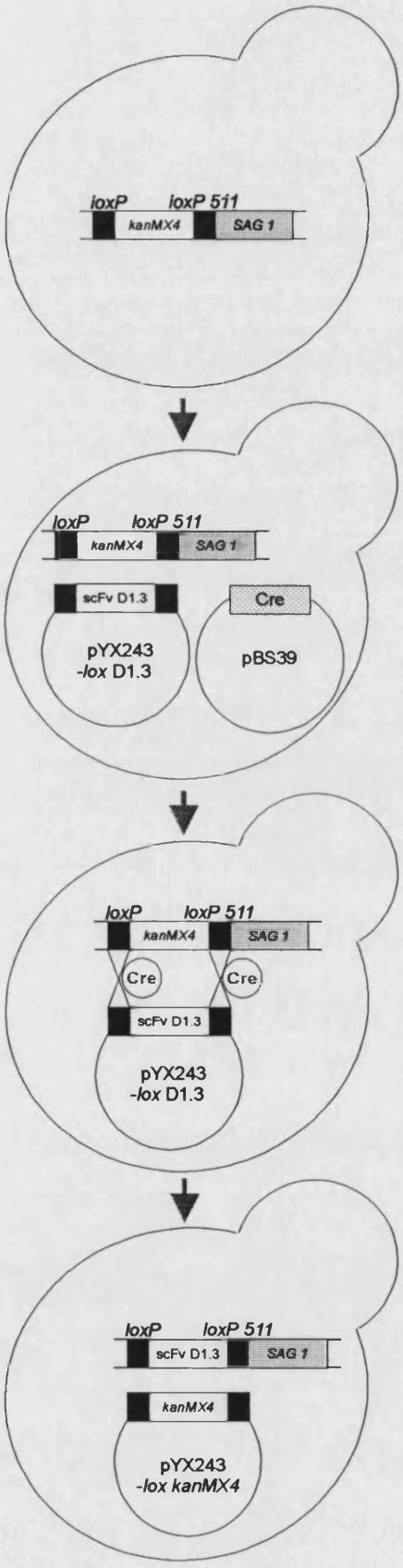
therefore solves the problem of gene conversion associated with the use of yeast marker genes (Güldener *et al.*, 1996).

The second stage in the creation of the agglutinin mutants was the construction of plasmids with the scFv D1.3 or HEWL open-reading frames flanked by the *loxP* and *loxP 511* sites. The *loxP/loxP 511* sites on these vectors were equivalent in orientation to the *loxP/loxP 511* sites either side of the *kanMX4* selectable marker in the genome. Special consideration in the design of the oligonucleotides used at this stage was also taken to conserve the open-reading frames of the agglutinin-fusions, especially avoiding any undesirable start or stop codons.

After co-transformation of one of the above vectors and the Cre recombinase expression vector, site-specific recombination between the *loxP* sites may be induced with galactose. Due to the fact that identity is required between *loxP* sites for recombination to take place (Table 5.1), exchange should only occur between the *kanMX4* cassette on the genome and the antibody (or antigen) DNA on the vector. After Cre recombinase expression, in-frame gene fusions with the relevant agglutinins should be created.

This Chapter describes the construction of both the agglutinin gene fusions. Also, the creation of the HEWL/α-agglutinin and scFv D1.3/α-agglutinin expression vectors is detailed.

Figure 5.3 - Schematic representation of the construction of the scFv D1.3-α-agglutinin gene fusion.



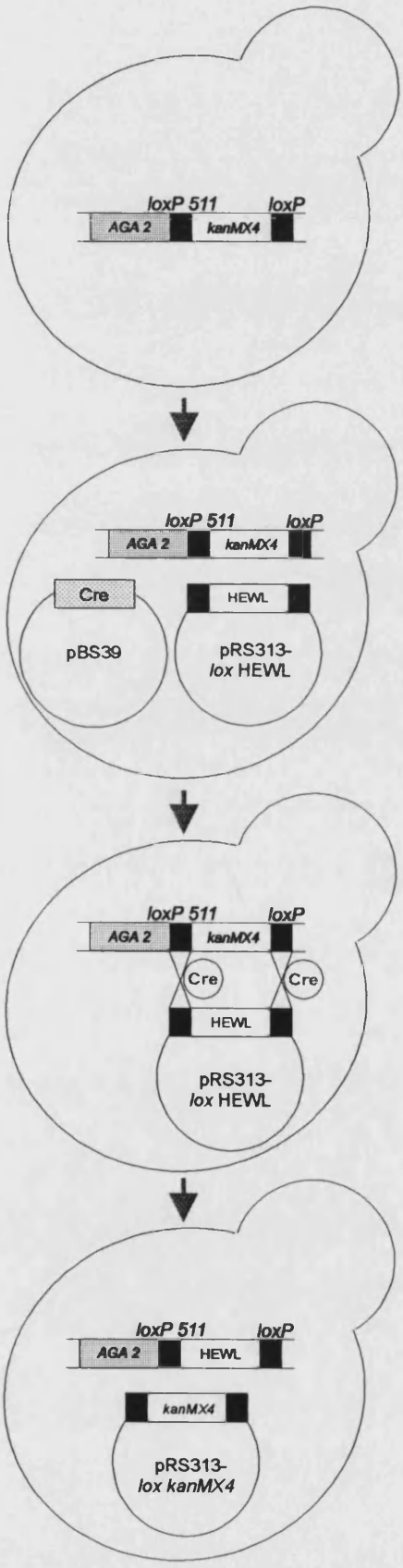
1) Integration of the PCR generated disruption cassette into the 5' terminal end of the α-agglutinin locus of α–cells.

2) Transformation of the Cre recombinase vector (pBS39) and exchange plasmid, pYX243-*lox* D1.3 into the disrupted α-strain.

3) Induction of Cre recombinase with galactose catalyses the intermolecular recombination between the *loxP* and *loxP 511* sites.

4) The desired scFv D1.3/α-agglutinin gene fusion was created.

Figure 5.4 A flow diagram summarising the production of the **a**-agglutinin-HEWL gene fusion.



1) Integration of the PCR generated disruption cassette into the 3' terminus of the **a**-agglutinin binding domain locus (*AGA2*) of **a**–cells.

2) Transformation of the Cre recombinase vector (pBS39) and exchange plasmid, pRS313-*lox* HEWL into the disrupted **a**-strain.

3) Induction of Cre recombinase with galactose catalyses the intermolecular recombination between the *loxP* and *loxP 511 sites.*

4) The desired **a**-agglutinin/ HEWL gene fusion was created.

5.2 MATERIALS

5.2.1 *Saccharomyces cerevisiae* strains

The two *S. cerevisiae* strains, FY22 and FY10 were isogenic derivatives of strain S288C. These strains were chosen because they were galactose inducible and contained the appropriate auxotrophic mutations for plasmid maintenance and for the proposed diploid antigen-selection strategy (Chapter 6). A diploid strain, FY1679 from the Fred Winston's collection was used as the source of DNA for the European Union yeast genome sequencing programme.

Table 5.2 *Saccharomyces cerevisiae* strains used.

Strain	Genotype	Source	Reference
FY10	*MATα ura3-52 leu2Δ1*	F. Winston, Harvard Medical School	Winston *et al.*, 1995
FY22	*MATa ura3-52 his3Δ200*	" "	" "

5.2.2 *Escherichia coli* strains and plasmids

The plasmids used in this chapter were either supplied as an *E. coli* slant culture (Table 5.3a) or in an aqueous DNA solution (Table 5.3b).

Table 5.3a Plasmids supplied in *E. coli* strains.

Plasmid	Strain	Genotype	Source	Reference
pBS39	DH1	*supE44 hsdR17 recA1 endA1 gyrA96 thi-1 relA1*	ATCC	Sauer, 1987
pUC8-HEWL	XL1-Blue	*supE44 hsdR17 recA1 endA1 gyrA96 thi-1 relA1 lac*$^-$ F'*[proAB*$^+$ *lacI*^{q} *lacZΔM15 Tn10(tet*^{r}*)]*	D.Jeenes, University of Norwich.	Jung *et al.*, 1980

When the plasmids were supplied in an aqueous DNA solution, the solution was transformed into TOP10F' *E. coli* cells (Section 2.2.15) and plasmid DNA was prepared (Section 2.2.17).

Table 5.3b Plasmids supplied in aqueous DNA solutions.

Plasmid	Source	Reference
pFA6a-*kanM4X*	A. E. Wheals, University of Bath.	Wach *et al.*, 1994
pUC119His6mycXba-scFv D1.3	R. E. Hawkins, University of Bristol.	Hawkins *et al.*, 1993
pTG5756	Transgene, France.	Langle-Rouault *et al.*, 1995
pRS313	S. Oliver, UMIST.	Sikorski and Hieter, 1989
pYX243	R+D Systems.	

For long term storage, glycerol stocks (Section 2.2.3) were made for all plasmids.

5.2.3 Plasmid descriptions

A) *The pUC8-HEWL vector*

The plasmid, pUC8-HEWL is an *E. coli* cloning vector and confers ampillicin resistance. It encodes the cDNA of HEWL which contains three silent polymorphisms: a GCT codon at Ala_{31}; an AAC codon at Asn_{106}; a GGC codon at Gly_{117} (D. J. Jeenes, personal communication) when compared to the published sequence (Jung *et al.*, 1980).

B) *The pUC119His6mycXba-scFv D1.3 vector*

The pUC119His6mycXba-scFv D1.3 vector contains the anti-HEWL V_H and V_L antibody domains joined together by a fifteen residue peptide linker, $(Gly_4Ser)_3$. It is a high copy number *E. coli* cloning vector containing a *c-myc* and His_6 tag which can be used for detection and purification, respectively. It confers ampicillin resistance and the vector promoter is from the *lac*Z gene,

therefore protein expression can be induced by isopropyl-β-D-thiogalactoside (IPTG).

C) *The Cre recombinase expression vector, pBS39 (Sauer,1987).*

The pBS39 vector (Figure 5.5) comprises of the Cre recombinase gene from bacteriophage P1, downstream from an inducible *GAL1* promoter. It contains a centromere (*CEN4*) and an autonomously replicating sequence (*ARS1*) to permit autonomous replication in *S. cerevisiae*. The ampillicin and *URA*3 auxotrophic markers were present for selection in *E. coli* and *S. cerevisiae*, respectively.

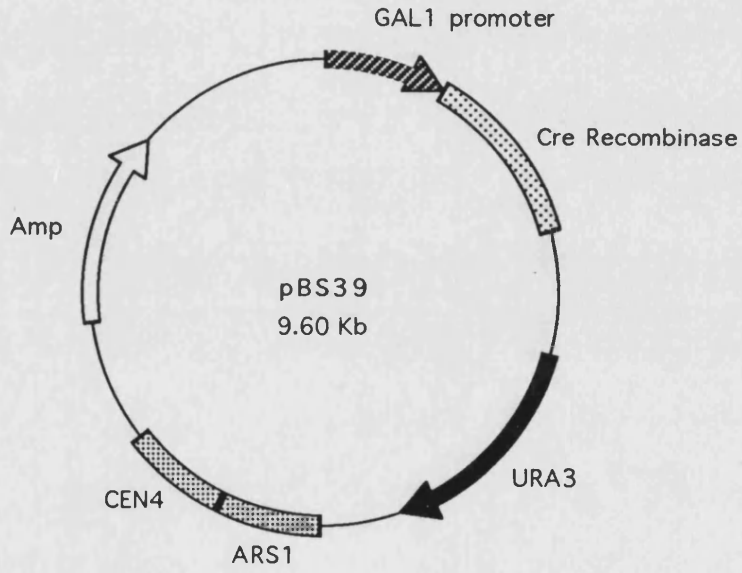


Figure 5.5 A plasmid map of the Cre recombinase expression vector, pBS39.

D) *The pFA6a-kanMX4 vector*

The pFA6a-*kanMX4* plasmid comprises of an ampicillin resistance selectable marker, an origin of replication for propagation in *E. coli* and a *kanMX4* marker cassette. The *kanMX4* cassette (Figure 5.6a) was cloned into the *Pme* I enzyme recognition site at the centre of the pFA6a multiple cloning site [MCS (Figure 5.6b)]. The flanking MCS sequences were essential in the design of the primers for the PCR-generated disruption cassettes due to the fact they contained the two 25 nucleotide stretches of homology essential for the 'splice by overlap extension' assembly reactions (Figures 5.8 and 5.13).

Two *Not* I enzyme recognition sites were present at the ends of the MCS for ease of excision of the *kan*MX4 cassette from the pFA6a plasmid.

Figure 5.6a A map of the *kanMX4* cassette. P_{TEF} and T_{TEF} indicate the promoter and terminator of the *A. gossypii* TEF gene, respectively, and *kan*^{r} denotes the open reading frame of the *E. coli* transposon, *Tn903*.

b- The sequence of pFA6a multiple cloning site (MCS). Important restriction enzyme sites are in bold and the primer recognition sites for the LFH-PCR assemblies are underlined.

Not I ← 5' *kanMX4* primer site

5' **GCGGCCGC**CAGCTGAAGCTTCGTACGCTGCAGGTCGACGGATCCCCGGGT
3' **CGCCGGCG**GTCGACTTCGAAGCATGCGACGTCCAGCTGCCTAGGGGCCCA

Pme I 3' *kanMX4* primer site →

5' TAATTAAGGCGCGCCAGATCT**GTTTAAAC**GAGCTCGAATTCATCGATGATAT
3' ATTAATTCCGCGCCGTCTAGA**CAAATTTG**CTCGAGCTTAAGTAGCTACTATA

Not I

5' CAGATCCACTAGTGGCCTAT**GCGGCCGC**
3' GTCTAGGTGATCACCGGATA**CGCCGGCG**

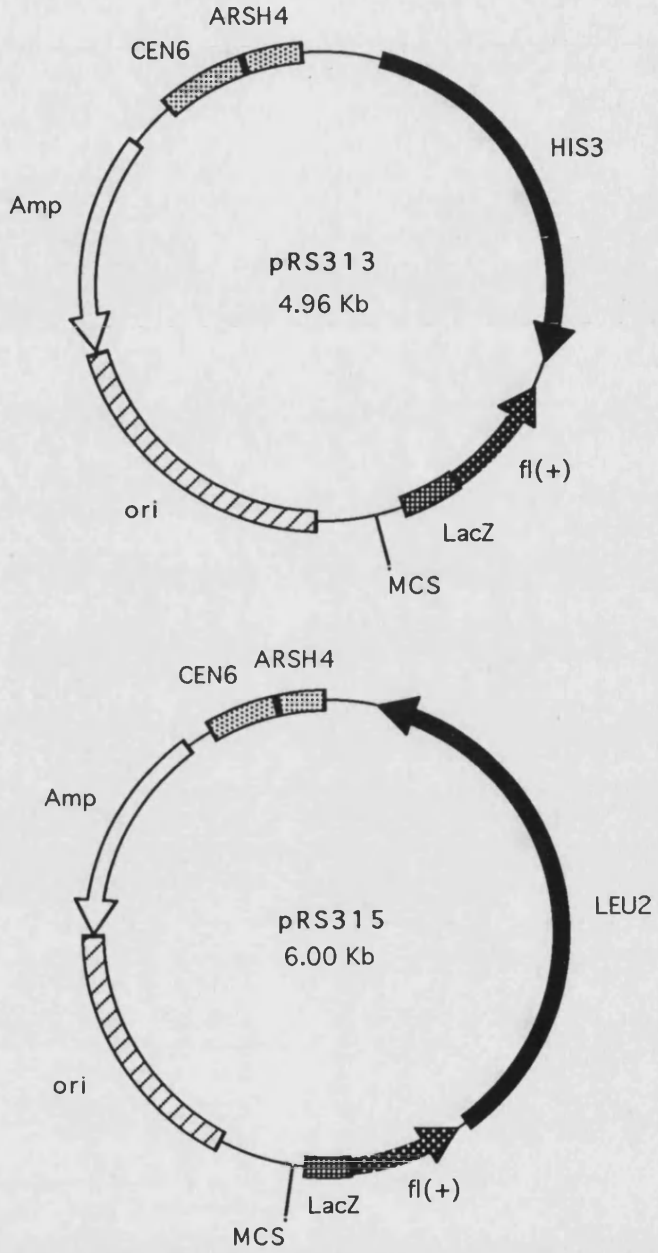
E) *The centromeric yeast shuttle vectors, pRS313 and pRS315*

The yeast shuttle vectors, pRS313 and pRS315 contain either a *HIS3* or *LEU2* selectable marker, respectively. They both have similar general structures (Figure 5.7a) and possess all the attributes of pBLUESCRIPT (Stratagene) including an extensive multiple cloning site (Figure 5.7b). Consequently, due to the fact they contain a bacterial origin of replication, all the standard DNA manipulations can be performed in *E. coli.* The *Not* I and *Apa* I enzyme recognition sites were found at either end of the multiple cloning site.

F) *The pYX243 vector*

The pYX243 plasmid is described in Section 3.2.3b (Figure 3.5).

Figure 5.7a Plasmid maps of the centromeric yeast shuttle vectors, pRS313 and pRS315.



b - Multiple cloning site for the pRS3XX vector series. All the vectors in the series have the same multiple cloning site, but not all the restriction sites are unique for each vector. The restriction sites *Not* I and *Apa* I are shown in bold.

Sac II *Not* I

Sac I *BstX* I *Eag* I *Xba* I *Spe* I *Bam*H I *Sma* I *Pst* I

TGGAGCTCCACCGCGGTG**GCGGCCGC**TCTAGAACTAGTGGATCCCCCGGGCTGCAG

*Bsp*D I *Kpn* I

*Eco*R I *Eco*R V *Hind* III *Cla* I *Sal* I *Xho* I *Apa* I *Acc*65 I

GAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGGG**GGGCCC**GGTACC

5.3 METHODS - CONSTRUCTION OF THE scFv D1.3-*SAG1* FUSION IN α-CELLS

5.3.1 Generation of the *loxP-kanMX4-loxP 511 SAG1* gene disruption cassette.

A) *Primer design*

The sequences of the four primers used to generate the *SAG1* disruption cassette (Table 5.4) were derived from the promoter and C-terminal regions of the α-agglutinin gene (*SAG1* locus). The DNA sequences were obtained from the *Saccharomyces cerevisiae* genome database [SGD (http://genome-www.stanford.edu/VL-yeast.html.)] and the accession number for the *SAG1* open reading frame is YJR004C.

The fundamental design parameters for primers were as described in section 2.2.4. The sagP1x and sagC1x primers were designed to have their 5' ends homologous to 25 nucleotides of the pFA6a-MCS, each side of the *kanMX4* cassette (Section 5.2.3d). Therefore, in the second PCR step, the disruption cassette can be assembled using a 'splicing by overlap extension' reaction (Figure 5.8). Both these primers also incorporated a *loxP* (or *loxP 511*) site and contained 20 nucleotides of homology to the relevant regions of the *SAG1* locus. The design of primer sagC1x required extra consideration because the *loxP 511* site is very adenine/thymine rich. To avoid any stop codons between the scFv D1.3 and C-terminal α-agglutinin gene fusion, two extra nucleotides (cytosine/adenine) were added after the *loxP 511* site, to maintain the open reading frame (Figure 5.9). The *loxP* and *loxP 511* sites were also 'inverted' due to the fact that only one reading frame was available which did not contain any stop codons.

B) *PCR amplification of the promoter and C-terminal α-agglutinin regions*

Two separate Expand™ PCRs (Section 2.2.7c) were performed to amplify the promoter and C-terminal regions of the α-agglutinin gene from yeast genomic DNA. The genomic DNA was prepared from the *S. cerevisiae* α-strain, FY10 following the procedure outlined in Section 2.2.21.

The primer combination sagP1x/alp1(p) was used to amplify the promoter region of the α-agglutinin gene. The reaction was performed for 25 cycles using the following parameters: [94°C, 2min; (94°C, 15s; 55°C, 30s; 72°C, 55s) x 10; (94°C, 15s; 55°C, 30s; 72°C, 55s + 20s/cycle) x 15; 72°C, 5min].

The C-terminal region of α-agglutinin was amplified using the primer combination sagC1x/α(F) and the PCR conditions for this reaction were the same as above, except the annealing temperature was 52°C.

Both PCR products were purified (Section 2.2.9b) and quantified by agarose gel electrophoresis (Section 2.2.8).

Figure 5.8 Schematic representation of the two-step PCR synthesis of the *loxP-kanMX4-loxP 511 SAG1* disruption cassette. L represents the α-agglutinin leader sequence and the different shaded boxes illustrate the promoter, N-terminal and C-terminal parts of the α-agglutinin gene (Figure is not to scale).

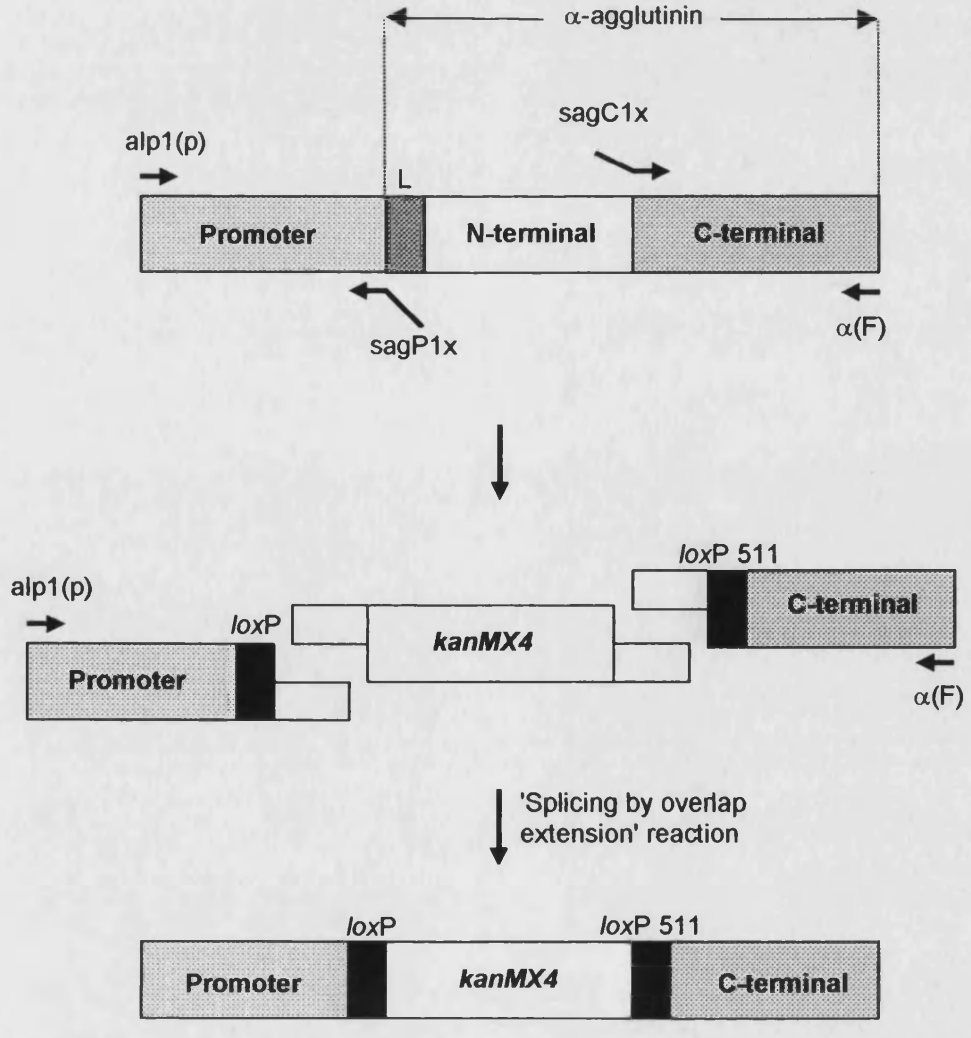


Figure 5.9 Schematic representation of the scFv D1.3/α-agglutinin gene fusion including sequence details of the regions encoding the N- and C-terminal ends. Restriction sites are represented in bold and the extra nucleotides are underlined.

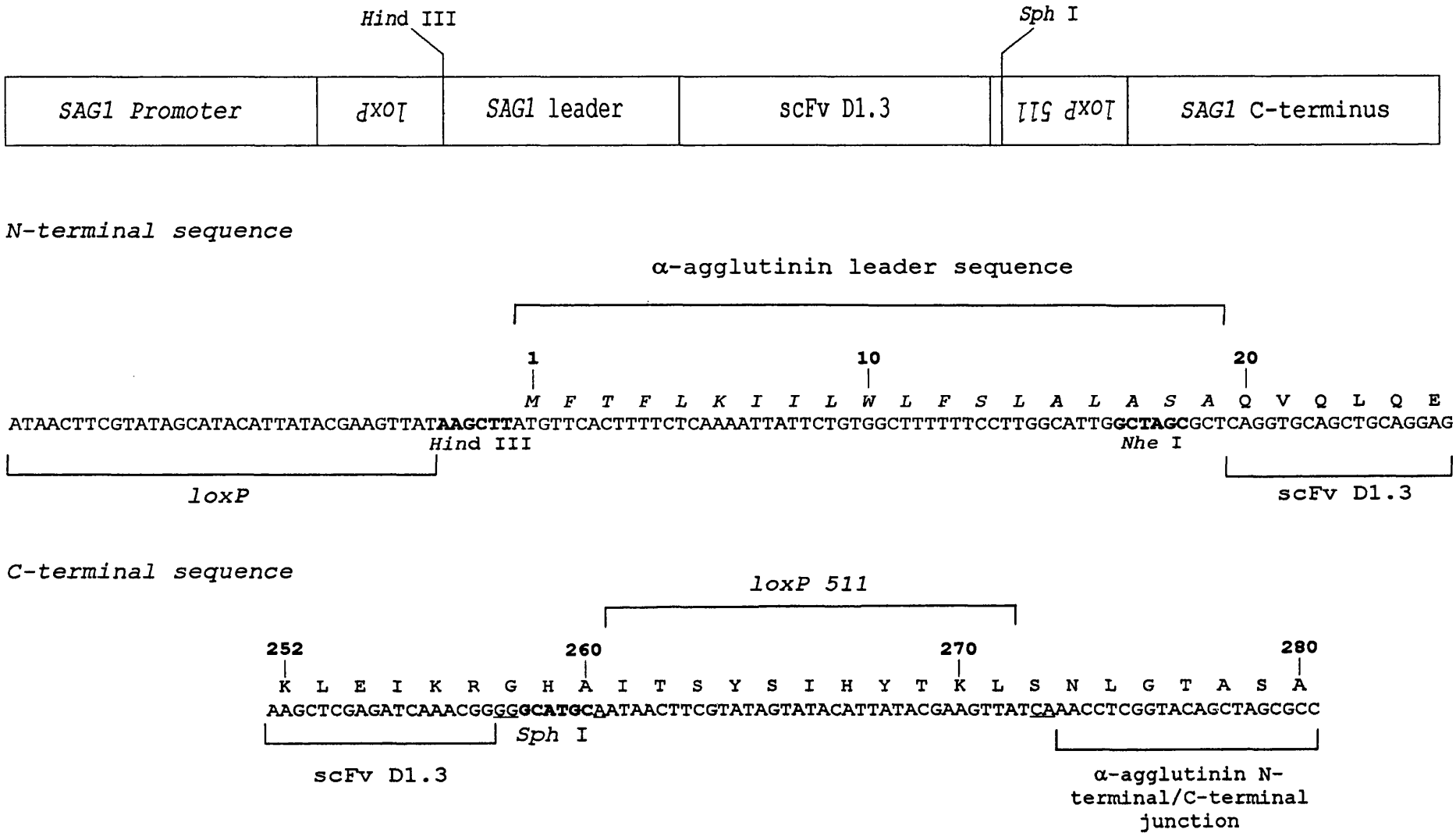


Table 5.4 Sequences of the oligonucleotides used in the generation of the *loxP-kanMX4-loxP 511 SAG1* disruption cassette. The *loxP* and *loxP 511* sites are underlined.

Primer Name	Sequence (5'-3')	Description (5'-3')
alp1(p)	TGATGATTTGAACGATCTCC	Binds 950-930 b.p. upstream from the *SAG1* start codon (Watson strand).
sagP1x	GGGGATCCGTCGACCTGCAG CGTACATAACTTCGTATAATG TATGCTATACGAAGTTATTTT GAAAACTGGCGAATTGT	25 n.t. of homology to the 5' end of pFA6a-MCS; 34 n.t. *loxP* site; 20 n.t. of homology to the promoter region directly before the *SAG1* start codon (Crick strand).
sagC1x	AACGAGCTCGAATTCATCGAT GATAATAACTTCGTATAGTAT ACATTATACGAAGTTATCAAA CCTCGGTACAGCTAGCGC	25 n.t. of homology to the 3' end of pFA6a-MCS; 34 n.t. *loxP 511* site; two extra n.t.; 20 n.t. of homology to the Watson strand of the N/C junction of α-agglutinin (975-955 b.p. upstream from the *SAG1* stop codon).
α(F)	AACTGTACAGTACCCGTT	Binds 18 n.t. directly after the *SAG1* stop codon (Crick strand).

C) *Digestion of pFA6a-kanMX4 vector with Not* I

The vector, pFA6a-*kanMX4* (~1μg) was digested with 10units of *Not* I overnight at 37°C (Section 2.2.12). The 1.5 k.b.p. *kanMX4* cassette was separated from the pFA6a plasmid by low melting point agarose gel electrophoresis (Section 2.2.8) and purified (Section 2.2.9b).

D) *PCR mediated generation of the SAG1 disruption cassette*

Approximately equal amounts (~50ng) of the *Not* I digested *kanMX4* cassette, the promoter and C-terminal α-agglutinin fragments were added together in an Expand™ PCR (Section 2.2.7c) with the primer pair alp1(p)/α(F). The reaction conditions were: [94°C, 2min; (94°C, 15s; 60°C, 2min; 68°C, 2min) x 10; (94°C, 15s; 60°C, 2min; 68°C, 2min + 20sec/cycle) x 15; 72°C, 5min]. The assembled product (3.6 k.b.p.) was separated from the other non-specific bands by low melting point agarose gel electrophoresis (Section 2.2.8) and recovered using β-agarase (Section 2.2.9a).

5.3.2 Integration of the *loxP-kanMX4-loxP 511 SAG1* disruption cassette into FY10

A) *Transformation of SAG1 gene disruption cassette into FY10*

The assembled *kanMX4 SAG1* disruption cassette (~250ng) was transformed into the *S. cerevisiae* α-strain (FY10) using a lithium acetate/polyethylene glycol transformation protocol (Section 2.2.24).

B) *Verification of G418^r transformants by PCR*

Yeast genomic DNA was made (Section 2.2.21) for three geneticin resistant transformants. Two sets of primers were used to check both novel joints that were formed upon integration (Table 5.5; Figure 5.10) via *Taq* PCRs (Section 2.2.7b).

Genomic integration of the disruption cassette at the promoter end of the *SAG1* locus was verified using the FL2B/Apl(seq) primers. The cycling parameters were: [(94°C, 1min; 50°C, 1min; 72°C, 1.30min) x 30; 72°C, 10min].

The C-terminal integration event was examined using the primer combination FL1F/YJR003C and the conditions were the same as above except that the elongation time was one minute. All PCR products were analysed by agarose gel electrophoresis (Section 2.2.8).

N.B. A wild type genomic DNA PCR control was run in conjunction.

C) *Analysis of the integration sites by DNA sequencing*

The primer pair APL(seq)/YJR003C was used to amplify a 4.1 k.b.p. PCR product containing both integration sites from the yeast genomic DNA of one positive transformant. An Expand™ PCR (Section 2.2.7c) was performed and the conditions were: [94°C, 2min; (94°C, 15s; 50°C, 30s; 68°C, 2.30min) x 10; (94°C, 15s; 50°C, 30s; 68°C, 2.30min + 20s/cycle) x 15; 72°C, 5min]. The total PCR (50μl) was resolved on a low melting point agarose gel (Section 2.2.8) and purified using β-agarase (Section 2.2.9a).

Five separate DNA sequencing reactions (Section 2.2.18) were carried out to check the various parts of the integrated construct. The primers Apl(seq) and YJR003C were used to analyse the promoter and C-terminal integration sites, respectively. The *loxP* and *loxP 511* sites were analysed with primers FL2B and FLIF. The middle section of the C-terminal α-agglutinin was examined with primer AggSeq2 (Figure 5.10). This α-strain containing the correct integrated *loxP-kanMX4-loxP 511* disruption cassette was named CAY1 (Table 5.11).

Figure 5.10 Diagrammatic representation of the primer sites used for the verification and sequencing of the integration events at the *SAG*1 locus. Symbols ' ✕ ' represent the novel joints formed by homologous recombination and ' → ' symbolises the oligonucleotides. Figure is not to scale.

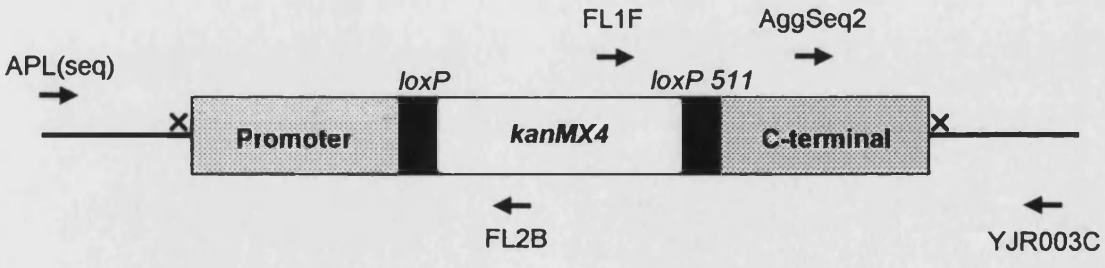
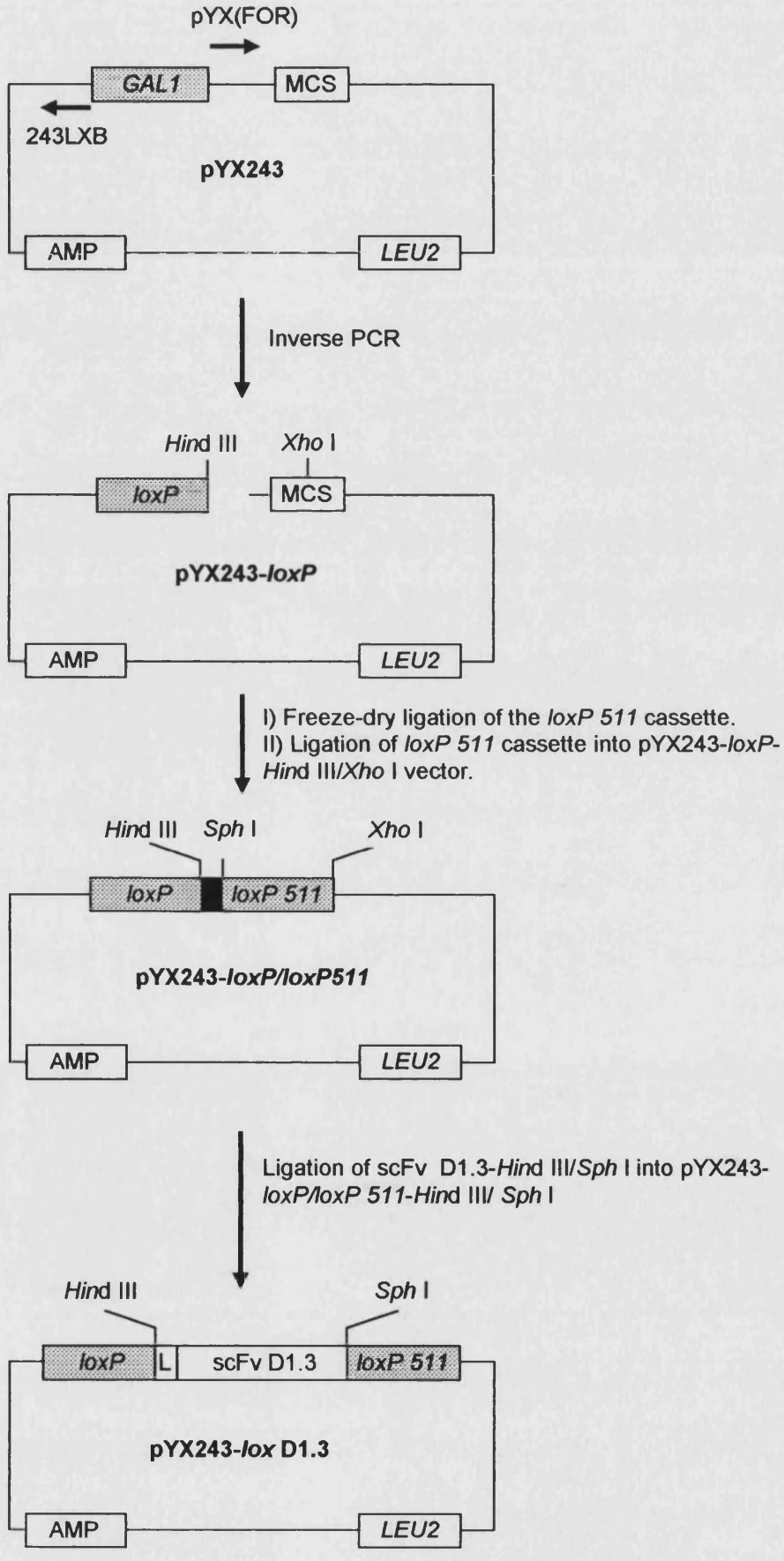


Table 5.5 Sequences of the oligonucleotides used in the verification and sequencing of the *SAG1* integration events.

Primer Name	Sequence (5'-3')	Description (5'-3')
AggSeq2	CCATCTGAAGAGCCCACTTT	434-454 b.p. upstream from the *SAG1* stop codon (Watson strand).
APL(seq)	TCGCCATAGTTTCCTAGAAC	1195-1215 b.p. from *SAG1* start codon (Watson strand).
FL2B	GGCGACAGTCACATCATGCC	163-183 b.p. downstream of 5' end of *kanMX4* (non-coding strand).
FL1F	CCTCGACATCATCTGCCC	162-180 b.p. upstream of 3' end of *kanMX4* (coding strand).
YJR003C	ATTATTGAATGGCGGTTAGC	337-357 b.p. downstream from the *SAG1* stop codon (Crick strand).

5.3.3 Construction of pYX243-*lox* D1.3 vector

Figure 5.11 A flow diagram summarising the construction of the vector, pYX243-*lox* D1.3.



A) *Construction of pYX243-loxP using inverse PCR*

An 8 k.b.p. inverse DNA fragment was generated using pYX243 (~75ng) as a DNA template and the primer pair pYX(FOR)/243LXB in an Expand™ PCR (Section 2.2.7c; Figure 5.11).The inverse PCR method requires the use of primers complementary to opposite strands of DNA and orientated so that primer extension proceeds 5' to 3' from each primer towards the other primer (Triglia *et al.,* 1988). The primer pYX(FOR) was designed to bind to the coding strand at the 3' end of the *GAL1* promoter and 243LXB was designed to bind to the non-coding strand at the 5' end of the *GAL1* promoter. The primer 243LXB also incorporated a *loxP* site and a *Hind* III restriction site. This vector created by inverse PCR did not have a *GAL1* promoter and was named pYX243-*loxP*. The basic design parameters for primers were followed (Section 2.2.4) and table 5.6 describes the oligonucleotides in more detail.

The cycling parameters for the PCR were: [94°C, 2min; (94°C, 15s; 50° C, 30s; 68°C, 6min) x 10; (94°C, 15s; 50°C, 30s; 68°C, 6min + 20s/cycle) x 15; 72°C, 6min]. Two identical PCR products (100μl) were analysed on a low melting point agarose gel (Section 2.2.8), purified (Section 2.2.9b) and the aliquots were pooled together for subsequent manipulations.

The vector, pYX243-*loxP* was digested with *Hind* III (10units) and *Xho* I (20units) overnight at 37°C (Section 2.2.12). This digested fragment was resolved on a low melting point agarose gel (Section 2.2.8) and purified (Section 2.2.9b).

B) *Freeze-dry ligation of loxP 511 cassette into pYX243-loxP*

A double stranded oligonucleotide *loxP 511* building block cassette (56 b.p.) was constructed by annealing oligonucleotides, loxP511A(NEW) and loxP511B(NEW) [Table 5.6]. The primers were designed to contain a *loxP 511* site, a nine nucleotide spacer region, a *Sph* I site and one extra nucleotide to maintain the open reading frame of the scFv D1.3-α-agglutinin fusion. They also both contained partial *Xho* I and *Hind* III recognition sites at appropriate ends for cloning into pYX243-*loxP*.

Each primer (200 pmoles) was phosphorylated at the 5' terminus with T4 polynucleotide kinase at 37°C for 60 minutes. The reaction mixture contained 10units of T4 polynucleotide kinase and 5µl of 10 x polynucleotide kinase buffer [700mM Tris-HCl (pH7.6), 100mM $MgCl_2$, 50mM DTT] in a total volume of 50 µl. After the incubation period, the T4 polynucleotide kinase was heat inactivated at 65°C for 20 minutes. Annealing was performed in a 0.5ml Eppendorf by mixing an equimolar amount of each oligonucleotide (0.1pmoles) and 2.5µl of T4 DNA ligase buffer (Section 2.2.13), in a final volume of 25µl. The lid of the 0.5ml Eppendorf was pierced with a needle and then the Eppendorf was dipped in liquid nitrogen for 60 seconds. The Eppendorf was then quickly transferred to a 500ml round bottom flask (pre-chilled at -20°C) and lyophilised using a Lyoprep-300 freeze-drier (International Equipment Company, Dunstable, UK) at less than 0.05mbar for 2 hours. The lyophilised pellet was resuspended in 25µl of water (4°C) and used as an insert in a standard ligation reaction (Section 2.2.13) with digested vector, pYX243-*loxP-Hind* III/*Xho* I. Ligations with molar ratios (insert:vector) of 50:1, 10:1 and 1:1 were carried out and transformed directly into *E. coli* cells (Section 2.2.15).

C) *Verification of loxP 511 site in vector pYX243-loxP/loxP 511*

Eight putative positive transformants were PCR screened (Section 2.2.16) with the primer pair InvpYX(FOR)/InvpYX(BACK) (Table 5.6). The PCR products were alcohol precipitated (Section 2.2.10) and the DNA pellets were resuspended in 10µl of water. Each precipitated PCR product was digested with *Sph* I (5units) overnight at 37°C (Section 2.2.12), in a total reaction volume of 15µl. The entire reactions were analysed by agarose gel electrophoresis (Section 2.2.8) and glycerol stocks (Section 2.2.3) were made of transformants that contained the vector with the *Sph* I recognition site. This constructed vector was then referred to as pYX243-*loxP/loxP 511*.

D) *PCR amplification of scFv D1.3*

The scFv D1.3 open reading frame was amplified by an Expand™ PCR (Section 2.2.7c) using the vector pUC119His6mycXba-scFv D1.3 as a DNA

template and the primer pair DF/DB(SphI) (Table 5.6). The primer DF was designed to anneal to the first 19 b.p. of the 5' end of scFv D1.3 and also encoded the α-agglutinin leader sequence. The oligonucleotide DB(SphI) annealed to the last 18 b.p. of scFv D1.3 (before the stop codon) and incorporated an extra two cytosine nucleotides to maintain the open reading frame of the scFv D1.3/α-agglutinin gene fusion. To facilitate cloning, either a *Hind* III site (DF) or *Sph* I site [DB(SphI)] were introduced to the 5' ends of the oligonucleotides. The PCR conditions for this amplification were: [94°C, 2min; (94°C, 15s; 60°C, 30s; 72°C, 50s) x 10; (94°C, 15s; 60°C, 30s; 72°C, 50s + 20s/cycle) x 15; 72°C, 5min]. The amplified DNA fragment (804 b.p.) was purified (Section 2.2.9b) and quantified using agarose gel electrophoresis (Section 2.2.8).

E) *Construction of plasmid pYX244-lox D1.3*

The vector, pYX243-*loxP/loxP 511* and the amplified scFv D1.3 fragment were digested with *Hind* III (10units) and *Sph* I (5units) overnight at 37°C (Section 2.2.12). Both were purified using the appropriate DNA extraction kits (Section 2.2.9b) and quantified by agarose gel electrophoresis (Section 2.2.8). ScFv D1.3-*Hind* III/*Sph* I was ligated with pYX243-*loxP/loxP 511-Hind* III/*Sph* I (Section 2.2.13) and transformed into *E. coli* cells (Section 2.2.15). The ampicillin-resistant colonies were PCR screened (Section 2.2.16) using primers InvpYX(FOR)/InvpYX(BACK). The PCR products were analysed by agarose gel electrophoresis (Section 2.2.8) and transformants with insert were grown overnight in 3ml of LB (Section 2.2.1). Glycerol stocks were prepared (Section 2.2.3) and the remaining cultures were used for plasmid DNA preparations (Section 2.2.17). DNA sequencing (Section 2.2.18) was performed using primers InvpYX(BACK) and InvpYX(FOR).

Table 5.6 Sequences of oligonucleotides used in the construction of vector pYX243-*lox* D1.3. The *loxP* and *loxP 511* sites are underlined and the restriction endonuclease recognition sites are in bold.

Primer Name	Sequence (5'-3')	Description (5'-3')
243LXB	GCGAG**AAGCTT**ATAACTTCGTA TAATGTATGCTATACGAAGTTAT AATGACAGCTCAGTTACA	5 n.t. overhang, a *Hind* III site, 34n.t. *loxP* site, 18n.t. of homology to non-coding strand of pYX243 (7934-7942).
pYX(FOR)	TACTTTAACGTCAAGGAG	18 n.t. of homology to the coding strand of pYX243 (8592-8610).
loxp511A(NEW)	**AGCTT**GCAGGTCAG**GCATGCA**A TAACTTCGTATAGTATACATTAT ACGAAGTTAT**C**	5 n.t. partial *Hind* III site; 9 n.t. spacer region; a *Sph* I site; 1 extra n.t.; 34 n.t. *loxP 511*; 1 n.t. partial *Xho* I.
loxp511B(NEW)	**TCGAG**ATAACTTCGTATAATGTA TACTATACGAAGTTAT**TGCATGC** CTGACCTGC**A**	5 n.t. partial *Xho* I site, 34 n.t. *loxP 511* site, 1 extra n.t.; a *Sph* I site; 9 n.t. spacer region,1 n.t. partial *Hind* III site.
InvpYX(FOR)	ATTAGCTCTACCACAGTG	18 n.t. of homology to the coding strand of pYX243 (7924-7942).
InvpYX(BACK)	CCAGTCACGACGTTGTAA	18 n.t. of homology to the non-coding strand of pYX243 (245-263).
DF	GCGAG**AAGCTT**ATGTTCACTTT TCTCAAAATTATTCTGTGGCTTT TTTCCTTGGCATTGGCTAGCGC TCAGGTGCAGCTGCAGGAGT	5 n.t. overhang; a *Hind* III site; 56 n.t. α-agglutinin leader sequence and 19 n.t. of homology to the 5' end of scFv D1.3 (coding strand).
DB(Sph1)	CTGAGATACT**GCATGC**CCCCGT TTGATCTCGAGCTT	10 n.t. overhang; a *Sph* I site; 2 cytosine n.t.; 18 n.t. of homology to the 3' end of scFv D1.3 (non-coding strand).

5.3.4 Verification of the Cre recombinase expression vector, pBS39

The Cre expression vector, pBS39 (~1μg) was digested with *Xho* I (20units) overnight at 37°C (Section 2.2.12) and analysed by agarose gel electrophoresis (Section 2.2.8). DNA sequencing was performed (Section 2.2.18) using primers Cre(Back) [Table 5.7] and pYX(FOR) [Table 5.6].

5.3.5 Transformation of the *SAG1* disrupted α-strain with pBS39 and pYX243-*lox* D1.3

The vectors, pYX243-*lox* D1.3 (~2μg) and pBS39 (~2μg) were co-transformed (Section 2.2.23) into the α-strain containing the *loxP-kanMX4-loxP 511* cassette integrated into its *SAG1* locus (CAY1; Table 5.11). Leu^+ Ura^+ transformants were PCR screened (Section 2.2.16) for the presence of both plasmids in separate PCRs. The pBS39 plasmid was detected using primers Cre(back)/pYX(FOR) and the pYX243-*lox* D1.3 plasmid was identified using primers InvpYX(FOR)/InvpYX(BACK). Each PCR was analysed by agarose gel electrophoresis (Section 2.2.8) and positive transformants were grown in 1ml of SC lacking uracil and leucine (supplemented with 2% raffinose) for 24 hours, at 30°C. Glycerol stocks were prepared (Section 2.2.3) and a 100μl aliquot of one culture was plated onto an SC agar plate containing 2% raffinose but lacking uracil and leucine. This was used as a stock plate.

Note: 2μg of pBS39 only, was also transformed into the *SAG1* disrupted α-strain (CAY1; Section 2.2.23) and SC without uracil/2% raffinose media was used throughout.

5.3.6 Solid plate galactose induction of Cre recombinase (Sauer *et al.*, 1987)

A single colony of the *SAG1* disrupted α-strain (CAY1; Table 5.11) containing plasmids pBS39 and pYX243-*lox* D1.3 was streaked onto a minimal media plate [0.67% (w/v) Bacto-yeast base without amino acids, 2% (w/v) Bacto-agar] supplemented with either 2% (w/v) glucose (SD) or 2% (w/v) galactose (SG). These plates were incubated at 30°C for 4 days. Three individual galactose grown colonies and one glucose grown colony were

picked and dispersed into 10ml of YPD. These cultures were incubated at 30°C, overnight. After this time, 200µl of each culture were transferred to fresh YPD (10ml) and incubated again overnight. This process was repeated for 3 days to remove both of the plasmids. The cultures were then counted using a haemocytometer, appropriately diluted in water and plated onto YPD agar plates. These plates were incubated for 16-24 hours at 30°C and the resulting colonies were replica-plated to YPD plates supplemented with 200mg/l geneticin (G418). These were then grown at 30°C for 16-24 hours and the geneticin-sensitive colonies were re-streaked onto YPD/G418 plates to determine 100% sensitivity. The colonies were also replica-plated onto the appropriate selective media to determine plasmid loss.

5.3.7 Liquid galactose induction of Cre recombinase

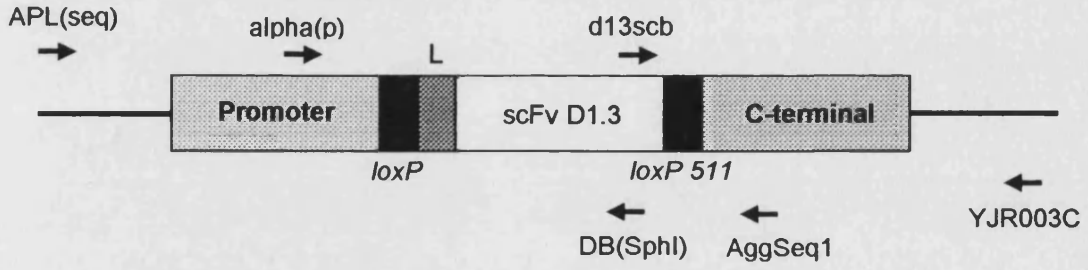
Three separate colonies of the *SAG1* disrupted α-strain (CAY1; Table 5.11) containing plasmids pBS39 and pYX243-*lox* D1.3 were grown in SC media without uracil and leucine [supplemented with 2% (w/v) raffinose] overnight at 30°C. The cells were then counted using a haemocytometer and diluted to 5 x 10^6 cells/ml in the same media. Galactose [20% (w/v)] or glucose [20% (w/v)] were added to a final concentration of 2% (w/v) and incubated at 30°C for either 5 or 24 hours. After this time, 200µl of the induced culture was transferred to 10ml of YPD and grown overnight at 30°C. This process was repeated for 3 days to remove the plasmids. The cells were then counted, diluted appropriately with water and spread onto YPD plates. These plates were incubated for 24-36 hours at 30°C and the resulting colonies were replica-plated onto YPD plates supplemented with 200mg/l geneticin (G418). These were then grown at 30°C for 16-24 hours and the geneticin-sensitive colonies were re-streaked onto YPD/G418 plates to determine 100% sensitivity. The colonies were also replica-plated onto the appropriate selective media to determine plasmid loss.

5.3.8 Verification of scFv D1.3 integration into the *SAG1* locus

Colonies that could not grow on geneticin (G418) were selected from the original YPD plate and transferred to 10 ml of YPD. These cultures were

grown overnight (30°C) and total yeast genomic DNA was prepared for each (Section 2.2.21). Individual genomic DNA samples were screened using three primer pairs in separate standard *Taq* PCRs (Section 2.2.7b). Primers Apl(seq)/DB(Sph1) analysed the 5' end, d13scb/YJR003C checked the 3' end and Apl(seq)/YJR003C amplified the total *SAG1* locus (Figure 5.12).The PCR conditions for the total amplification of the *SAG1* locus were: [(94°C, 1min; 50°C, 1min; 72°C, 3min30s) x 30; 72°C, 10min] and the PCR parameters for the amplification of the 5' and 3' ends were the same, except the elongation time was 2 minutes.

Figure 5.12 illustrates the primer sites used for the analysis and sequencing of scFv D1.3 in the *SAG1* locus. Symbol '→' represents the oligonucleotides (not to scale).



The PCR products were analysed by agarose gel electrophoresis (Section 2.2.8) and one positive integrant was chosen for DNA sequencing. An Expand™ PCR (Section 2.2.7c) was used to amplify the total scFv D1.3-*SAG1* gene fusion with primers APL(seq)/YJR003C. The PCR conditions were: [94°C, 2min; (94°C, 15s; 50°C, 30s; 68°C, 3min) x 10; (94°C, 15s; 50°C, 30s; 68°C, 3min + 20s/cycle) x 15; 72°C, 5min] and the 3.4 k.b.p. DNA fragment obtained was purified using β-agarase (Section 2.2.9a). The scFv D1.3, *loxP* and *loxP 511* sites were sequenced using primers AggSeq1 and alpha(p) [Table 5.7]. The α-strain containing the correctly integrated scFv D1.3 at the *SAG1* locus was named CAY2 (Table 5.11).

Table 5.7 Sequences of the oligonucleotides used in the screening of pBS39 and for the sequencing of integrated scFv D1.3 at the *SAG1* locus.

Primer Name	Sequence (5'-3')	Description (5'-3')
d13scb	CCAAGCTCGAGATCAAAC GG	Binds to the coding strand at the 3' end of scFv D1.3, before the *Sph* I site.
Cre(Back)	TTGAAGGCTCTCAAGGG C	Binds 4 b.p. after the *Sal* I site on pBS39 (non-coding strand).
Agg Seq 1	GCCTGTTTCTACTGTGGA AAT	Binds to the Crick strand of *SAG1*, 859-880 b.p. upstream from the *SAG1* stop codon.
alpha(p)	ATCGGCACCTCAATTCAA TG	Binds to the Watson strand of *SAG1*, 462-482 b.p. upstream from start.

5.3.9 Construction of plasmid pRS315-*lox* D1.3

A) *Removal of Kluyveromyces lactis URA3 marker from TG5756 to generate pRS315-Apa I/Not I*

A restriction digestion (Section 2.2.12) with enzymes *Apa* I and *Not* I was performed sequentially on plasmid TG5756 because the reaction requirements were incompatible. The vector, TG5756 (~1.5μg) was first digested with 10units of *Apa* I overnight at 25°C. The linearised plasmid was purified (Section 2.2.9b) and then incubated with *Not* I (10units) overnight at 37°C. An 6 k.b.p DNA fragment was resolved by low melting point agarose gel electrophoresis (Section 2.2.8) and recovered using the DNA Clean Up System (Section 2.2.9b).

B) *PCR amplification of loxP-scFv D1.3-loxP 511*

Primers pRS315(ApaI)/pRS315(NotI) were designed to anneal to the pYX243-*lox* D1.3 vector either side of the *loxP/loxP 511* sites (Table 5.8). To facilitate cloning, an *Apa* I or a *Not* I recognition site were incorporated into the 5' ends of the oligonucleotides. An Expand™ PCR (Section 2.2.7c) was used to amplify the 907 b.p. fragment containing the *loxP/loxP 511* sites and the scFv D1.3 gene. The following PCR conditions were used: [94°C, 2min; (94°C, 15s; 55°C, 30s; 72°C, 1min) x 10; (94°C, 15s; 55°C, 30s; 72°C, 1min + 20s/cycle) x 15; 72°C, 5min]. The PCR product was purified (Section 2.2.9b) and then digested sequentially with *Not* I and *Apa* I (Section 5.3.9a). The double digested product was then quantified using agarose gel electrophoresis (Section 2.2.8).

C) *Cloning of loxP-scFv D1.3-loxP 511 into pRS315*

The scFv D1.3-*Apa I/Not* I insert was ligated into plasmid pRS315-*Apa I/Not* I (Section 2.2.13). The ligation reaction was transformed into *E. coli* (Section 2.2.15) and ampicillin-resistant colonies were PCR screened (Section 2.2.16) to check for the scFv D1.3 insert using primers T7/M13 (Table 5.8). The PCR products were analysed by agarose gel electrophoresis (Section 2.2.8) and positive transformants were grown overnight in 3ml of LB (Section 2.2.1). Glycerol stocks (Section 2.2.3) and plasmid DNA (Section 2.2.17) were prepared. The construct of one transformant was checked for mutations using primers T7 and M13 by DNA sequencing (Section 2.2.18).

5.3.10 Transformation of the *SAG1* disrupted α-strain with pBS39 and pRS315-*lox* D1.3

The vectors, pRS315-*lox* D1.3 (~2μg) and pBS39 (~2μg) were co-transformed (Section 2.2.23) into α-strain containing the *loxP-kanMX4-loxP 511* cassette integrated into its *SAG1* locus (CAY1; Table 5.11). Ura^+ Leu^+ transformants were PCR screened (Section 2.2.16) for both plasmids using the primer pairs T7/M13 and pYX(FOR)/Cre(Back) for detection of pRS315-*lox* D1.3 and pBS39, respectively. Each PCR was analysed by agarose gel electrophoresis (Section 2.2.8) and positive transformants were grown in 1ml

of SC lacking uracil and leucine [supplemented with 2% (w/v) raffinose] for 24 hours (30°C). Glycerol stocks were made (Section 2.2.3) and a 100µl aliquot of one culture was plated onto an SC without uracil and leucine containing 2% (w/v) raffinose] agar plate. This was used as a stock plate.

Table 5.8 Sequences of the oligonucleotides used in the construction and sequencing of plasmid pRS315-*lox* scFv D1.3. Enzyme recognition sites are in bold.

Primer name	Sequence (5'-3')	Description (5'-3')
pRS315(ApaI)	CAGAT**GGGCCC**TGTAACT GAGCTGTCATT	5 n.t. overhang: an *Apa* I site; 18 n.t. of homology to the coding strand of pYX243-*lox* D1.3 directly before the *loxP* site.
pRS315(NotI)	CTGAGATACT**GCGGCCGC** TGAGCTCGAGATAACTTC	10 n.t. overhang : a *Not* I site; 18 n.t. of homology to the non-coding strand of pYX243-*lox* D1.3 (primer annealed to 8 n.t. of the *loxP 511* site and 10 n.t. to the following vector sequence).
T7	GTAATACGACTCACTATAG GGC	Binds 13 b.p. downstream from the *Apa* I site on pRS315 (non-coding strand).
M13	GGAAACAGCTATGACCAT G	Binds 67 b.p. upstream from the *Not* I site on pRS315 (coding strand).

5.4 METHODS - CONSTRUCTION OF THE *AGA2*-HEWL FUSION IN a-CELLS

The methods described in this section are similar to the construction of the scFv D1.3-*SAG1* gene fusion in α-cells.

5.4.1 Generation of the *loxP 511-kanMX4-loxP AGA2* gene disruption cassette.

A) *Primer design*

The sequences of the four primers used to generate the *AGA2* disruption cassette were derived from the open-reading frame of the **a**-agglutinin gene (*AGA2* locus) and the surrounding loci. The DNA sequences were obtained from the *Saccharomyces cerevisiae* genome database (SGD) and the accession number for the *AGA2* open reading frame was YGL032C.

Figure 5.13 outlines the scheme for the production of the *loxP 511-kanMX4-loxP AGA2* disruption cassette. A description of each primer can be found in table 5.9 and the basic design parameters for primers were as described in Section 2.2.4. The AGA2FUSP(BACK) and HOP2P(FOR) were designed to have their 5' ends homologous to 25 nucleotides of the pFA6a-MCS, each side of the *kanMX4* cassette (Section 5.3.2d). These two primers also incorporated the *loxP* and *loxP 511* sites. The primer, AGA2FUSP(BACK) was designed to remove the stop codon at the 3' end of the *AGA2* ORF and the *loxP 511* sequence was kept in a frame which did not encode any stop codons (Figure 5.14). The primer, HOP2(BACK) was designed to anneal directly after the *HOP2* gene to avoid disrupting its function.

Figure 5.13 Schematic representation of the two-step PCR synthesis of the *loxP 511-kanMX4-loxP AGA2* disruption cassette. RPL30A is a ribsomal protein present 219 b.p. upstream from the *AGA2* locus. *HOP2* is a meiosis-specific gene important in the process of pairing homologous chromosomes and is 634b.p. downstream from the *AGA2* locus. (Figure is not to scale).

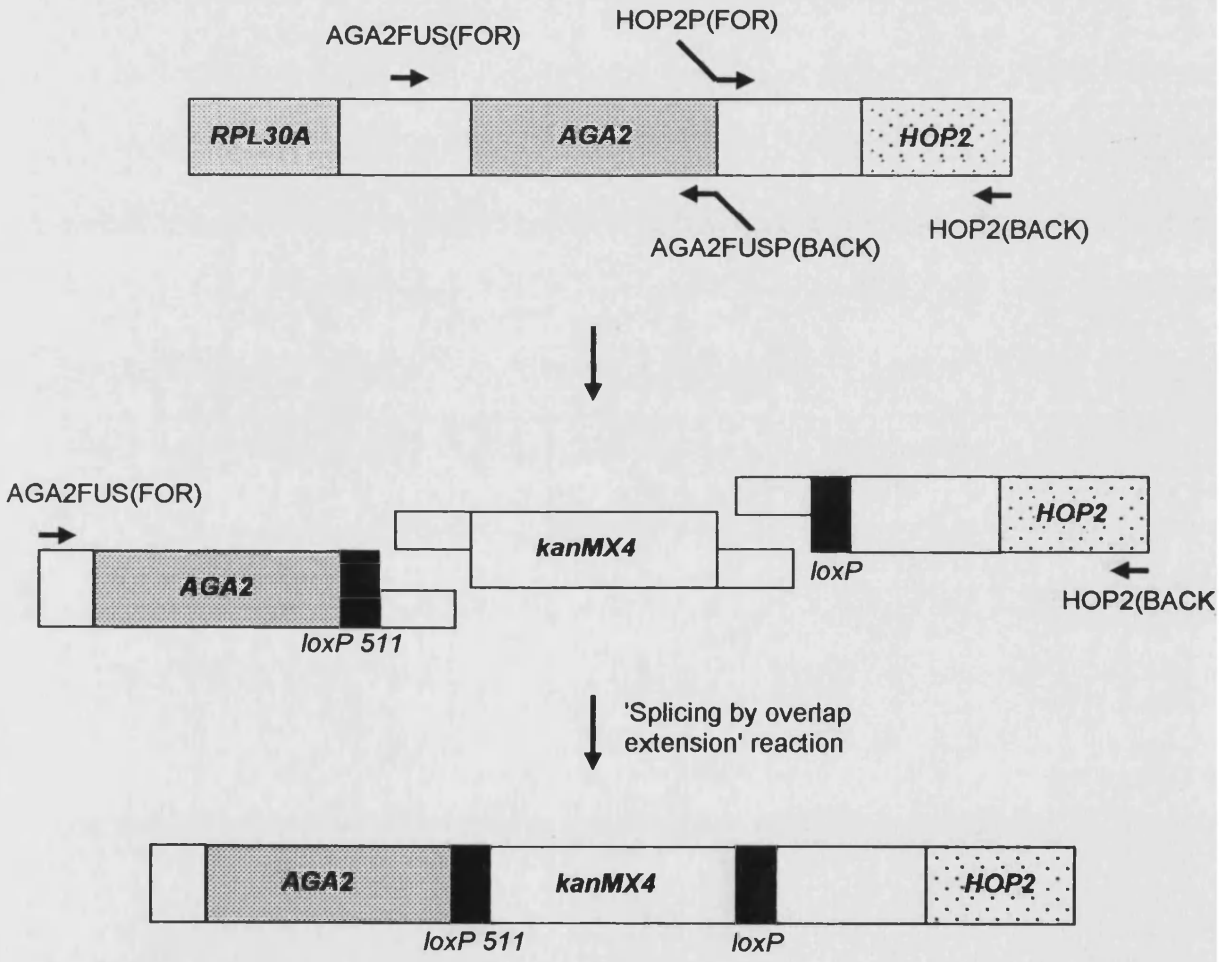
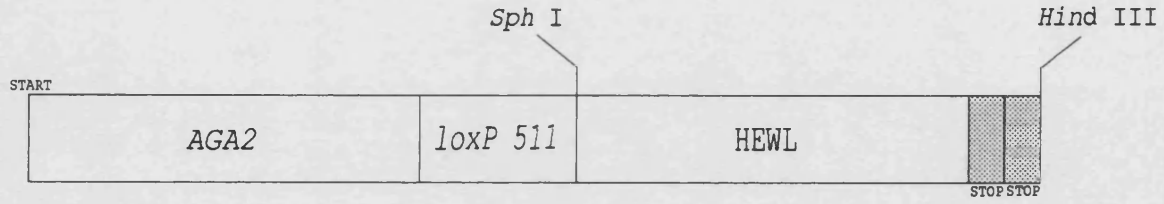
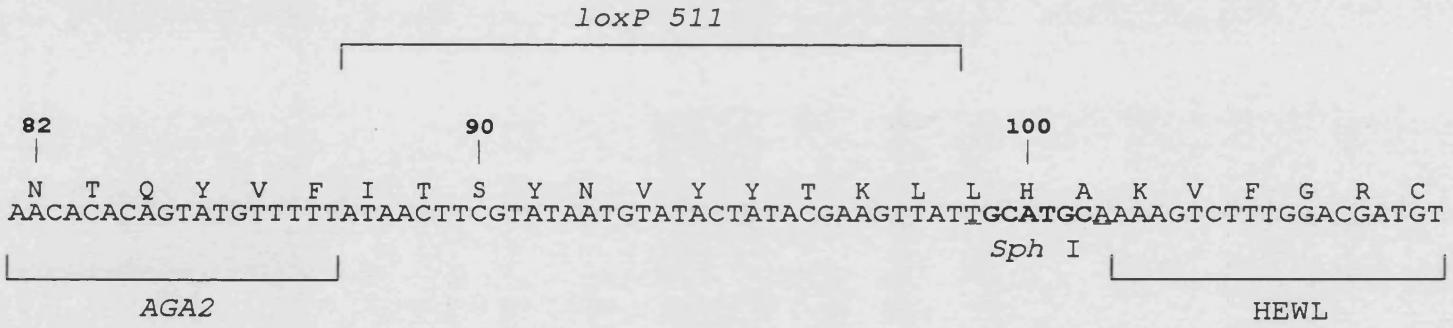


Figure 5.14 Schematic representation of the **a**-agglutinin/HEWL gene fusion including details of the 5' and 3' recombination sites. Restriction sites are represented in bold and the extra nucleotides are underlined. The stop codons are shown using the letter O.



5' DNA recombination site



3' DNA recombination site

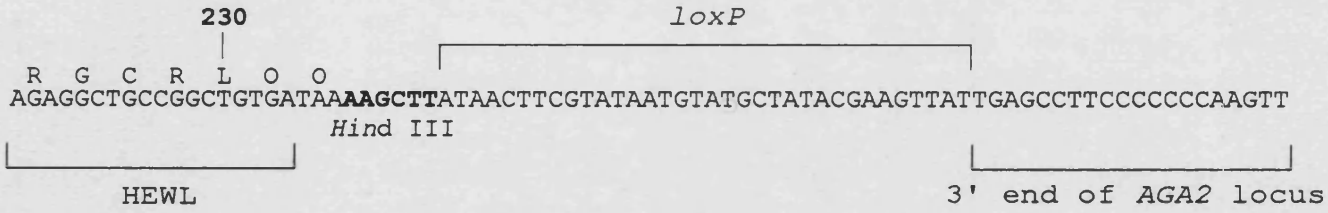


Table 5.9 Sequences of the oligonucleotides used in the generation of the *loxP 511-kanMX4-loxP AGA2* disruption cassette. The *loxP* and *loxP 511* sites are underlined.

Primer name	Sequence (5'-3')	Description (5'-3')
AGA2FUS(FOR)	TCGAACTTATTATGTACAAT	Binds 604-624 b.p. upstream from the *AGA2* start codon (Watson strand).
AGA2FUSP(BACK)	GGGGATCCGTCGACCTGCAG CGTACATAACTTCGTATAGTAT ACATTATACGAAGTTATAAAAA CATACTGTGTGTTTA	25 n.t. of homology to the 5' end of pFA6a-MCS; a 34 n.t. *loxP 511* site; 20 n.t. of homology to the *AGA2* gene directly upstream from the stop codon (Crick strand).
HOP2P(FOR)	AACGAGCTCGAATTCATCGAT GATAATAACTTCGTATAATGTA TGCTATACGAAGTTATTGAGCC TTCCCCCCCAAGTT	25 n.t. of homology to the 3' end of pFA6a-MCS; a 34 n.t. loxP site; 20 n.t. of homology to the *AGA2* locus directly from the stop codon (Watson strand).
HOP2(BACK)	ATTAACTTGTAGTCGGATCG	Binds 1323-1343 b.p. downstream from the *AGA2* stop codon (Crick strand).

B) *PCR amplification of the AGA2 and HOP2 gene fragments*

Two separate Expand™ PCRs (Section 2.2.7c) were performed to generate the *AGA2* and *HOP2* regions of the *AGA2* disruption cassette from FY22 genomic DNA (Section 2.2.21).

The primer combination AGA2FUS(FOR)/ AGA2FUSP(BACK) was used to amplify the *AGA2* DNA fragment (884 b.p.) and the reaction

conditions were: [94°C, 2min; (94°C, 15s; 50°C, 30s; 72°C, 1min) x 10; (94°C, 15s; 50°C, 30s; 72°C, 1min + 20s/cycle) x 15; 72°C, 5min].

The *HOP2* DNA fragment (1343 b.p.) was amplified using the primer pair HOP2P(FOR)/HOP2(BACK) and the cycling parameters were the same as above, except the annealing temperature was 45°C.

Both PCR products were purified (Section 2.2.9b) and quantified by agarose gel electrophoresis (Section 2.2.8).

C) *PCR-mediated generation of the loxP 511-kanMX4-loxP AGA2 disruption cassette*

Approximately equal amounts (~50ng) of the *Not* I digested *kanMX4* cassette, the *AGA2* and *HOP2* PCR fragments were added together in a standard Expand™ PCR (Section 2.2.7c) with the primer pair AGA2FUS(FOR)/HOP2(BACK). The reaction conditions were: [94°C, 2min; (94°C, 15s; 50°C, 2min; 68°C, 2min30s) x 10; (94°C, 15s; 50°C, 2min; 68°C, 2min30s+ 20s/cycle) x 15; 72°C, 5min].

The assembled product (3.8 k.b.p.) was separated from the other non-specific bands by low melting point electrophoresis (Section 2.2.8) and purified using β-agarase (Section 2.2.9a).

5.4.2 Integration of the *loxP 511-kanMX4-loxP AGA2* disruption cassette into the FY22 genome

A) *Transformation of the AGA2 gene disruption cassette into FY22*

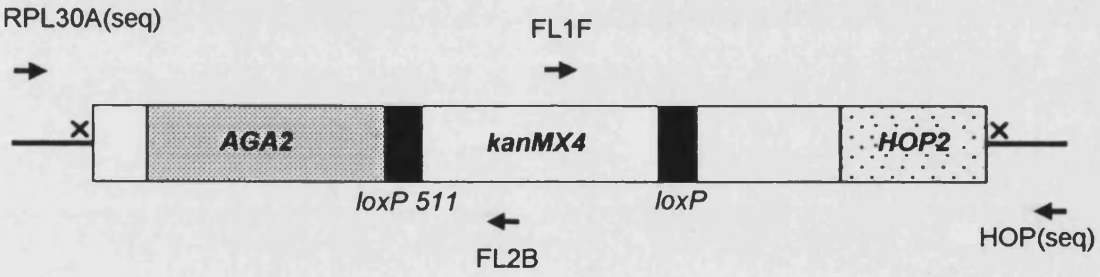
Approximately 250ng of the assembled *kanMX4 AGA2* disruption cassette was transformed into the *S. cerevisiae* **a**-strain (FY22) using a lithium acetate/polyethylene glycol transformation protocol (Section 2.2.24).

B) *Verification of G418^r transformants by PCR*

Genomic DNA was prepared (Section 2.2.21) for two geneticin resistant transformants. Two sets of primers (Tables 5.5 and 5.10) were used to check the two novel joints formed upon homologous recombination by standard *Taq* PCRs (Section 2.2.7b).

Integration of the disruption cassette at the 5' end of the *AGA2* locus was verified using primers RPL30A(seq)/FL2B (Tables 5.10 and 5.5). The 3' *HOP2* integration event was verified using primers FLIF/HOP2(seq) [Tables 5.5 and 5.10]. Figure 5.15 illustrates the regions of the *AGA2* locus that were amplified by the designed primers. The reaction conditions were the same for both PCRs: [(94°C, 1min; 50°C, 1min; 72°C, 1min) x 25; 65°C, 5min] and the PCR products were analysed by agarose gel electrophoresis (Section 2.2.8).

Figure 5.15 Diagrammatic representation of the primer sites used for the verification and sequencing of the integration events at the *AGA2* locus. Symbols '✕' represents the novel joints formed by homologous recombination and ' → ' symbolises the oligonucleotides (Figure is not to scale).



C) *Analysis of the AGA2 integration sites by DNA sequencing*

The primer pair RPL30A(seq)/HOP(seq) [Table 5.10] was used to amplify an 4.0 k.b.p. PCR product containing both the integration sites from the genomic DNA of one of the positive transformants. An Expand ™ PCR was followed (Section 2.2.7c) using the reaction conditions: [94°C, 2min; (94°C, 15s; 50°C, 30s; 68°C, 3min) x 10; (94°C, 15s; 50°C, 30s; 68°C, 3min+ 20s/cycle) x 15; 72°C, 5min]. The total PCR (50μl) was loaded onto a low melting point agarose gel (Section 2.2.8) and recovered using β-agarase (Section 2.2.9a). Two separate DNA sequencing reactions (Section 2.2.18) were carried out with the primers FL1F and FL2B to check the *loxP/loxP 511* sites of the integrated construct. The constructed **a**-strain containing the correctly integrated *loxP 511-kanMX4-loxP* disruption cassette at its *AGA2* locus was named CAY3 (Table 5.11).

5.4.3 Construction of the pRS313-*lox* HEWL plasmid

A) *Cloning of HEWL into pYX243-loxP/loxP 511*

The HEWL gene was amplified using pUC8-HEWL as a DNA template and the primer pair, LysF(SphI)/LysB(HindIII), in an Expand™ PCR (Section 2.2.7c). The primer, LysF(SphI) was designed to anneal to the first 18 b.p. of the 5' end of HEWL and incorporated an extra adenine nucleotide to keep the 3' terminal *AGA2* gene fusion with HEWL in the correct ORF. The oligonulceotide, LysB(HindIII) annealed to the last 15 b.p. of the 3' end of HEWL and also encoded two extra stop codons (Figure 5.14). To facilitate cloning, either a *Sph* I [LysF(SphI)] or a *Hind* III [Lys(HindIII)] site was introduced to the 5' ends of the oligonucleotides (Table 5.10). The reaction conditions were: [94°C, 2min; (94°C, 15s; 50°C, 30s; 72°C, 1min) x 7; (94°C, 15s; 50°C, 30s; 72°C, 1min+20s/cycle) x 8; 72°C, 5min]. The PCR (472 b.p.) was analysed by agarose gel electrophoresis (Section 2.2.8) and purified (Section 2.2.9b).

The vector, pYX243-*loxP/loxP 511* and the amplified HEWL fragment were digested with *Hind* III (10units) and *Sph* I (5units) overnight at 37°C (Section 2.2.8). Buffers were removed using the appropriate DNA extraction kits (Section 2.2.9b) and the digested fragments were quantified by agarose gel electrophoresis (Section 2.2.8). These were then ligated together (Section 2.2.13) and directly transformed into *E. coli* cells (Section 2.2.15). Ampicillin-resistant colonies were PCR screened (Section 2.2.16) using primers InvpYX(BACK)/InvpYX(FOR) and positive transformants were grown overnight in 3ml of LB (Section 2.2.1). Glycerol stocks (Section 2.2.3) and 'minipreps' (Section 2.2.17) were prepared from the overnight cultures. DNA sequencing was performed (Section 2.2.18) on one positive transformant using primers InvpYX(FOR) and InvpYX(BACK).

B) *Sub-cloning of loxP 511-HEWL-loxP into vector, pRS313*

The *loxP 511-HEWL-loxP* DNA fragment was amplified from construct pYX243-*lox* HEWL using the primer pair pRS315(ApaI)/pRS315(NotI) [Table 5.8]. An Expand™ PCR (Section 2.2.7c) was performed using the following conditions: [94°C, 2min; (94°C, 15s; 55°C, 30s; 72°C, 1min) x 7; (94°C, 15s;

55 °C, 30s; 72°C, 1min+20s/cycle) x 8; 72°C, 5min]. The PCR product was examined by agarose gel electrophoresis (Section 2.2.8) and the PCR buffers were removed (Section 2.2.9b).

The vector, pRS313 and the amplified *loxP 511*-HEWL-*loxP* fragment were digested sequentially with *Apa* I/*Not* I (Section 5.3.9a) and ligated together (Section 2.2.13). After transformation (Section 2.2.15), the ampicillin-resistant transformants were PCR screened using primers T7 and M13 (Section 2.2.16). Glycerol stocks (Section 2.2.3) and 'minipreps' (Section 2.2.17) were prepared from positive colonies which had been grown overnight in 3ml of LB (Section 2.2.1). Primers T7 and M13 were used to sequence one positive transformant (Section 2.2.18).

Table 5.10 Sequences of the oligonucleotides used in the verification of the *AGA2* integration events and the PCR amplification of HEWL. Enzyme recognition sites are in bold.

Primer Name	Sequences (5'-3')	Description (5'-3')
RPL30A(seq)	GGCTAAGTCTGCTGGTACTC	Binds 701-721 b.p. upstream from the *AGA2* start codon (Watson strand).
HOP2(seq)	AATAACTTGTAGTCGGATCG	Binds 1426-1446 b.p. downstream from the *AGA2* stop codon (Crick strand).
LysF(SphI)	CTGAGATACT**GCATGC**AAAAGTCTTTGGACGATGT	10 n.t. overhang; a *Sph* I site; one adenine n.t.; 18 n.t. of homology to the 5' end of HEWL (coding strand).
LysB(HindIII)	GCGAG**AAGCTT**TTATCACAGCCGGCAGCCTCT	5 n.t. overhang; a *Hind* III site; two stop codons; 15 n.t. of homology to the 3' end of HEWL (non-coding strand).

5.4.4 Transformation of FY22 with pBS39 and pRS313-*lox* HEWL

The vectors, pRS313-*lox* HEWL (~2 μg) and pBS39 (~2 μg) were co-transformed into the **a**-strain containing the *loxP 511-kanMX4-loxP* cassette integrated into its *AGA2* genome (CAY3; Section 2.2.23). His^+ Leu^+ transformants were PCR screened (Section 2.2.16) for the presence of both plasmids in separate PCRs. The plasmid pBS39 was detected using primers Cre(back)/pYX(FOR) and the pRS313-*lox* HEWL vector was identified using the primer combination T7/M13. Positive transformants were grown in 1ml of SC lacking uracil and histidine [supplemented with 2% (w/v) raffinose] for 24 hours (30°C). Glycerol stocks were prepared (Section 2.2.3) and a 100μl aliquot of one culture was plated onto an SC without uracil and histidine (containing 2% (w/v) raffinose) agar plate. This was used as a stock plate.

Note: pBS39 (~2μg) only was also transformed into the FY22 *AGA2* disrupted strain (Section 2.2.23) and SC without uracil/2% raffinose media was used throughout.

5.4.5 Cre-mediated integration of HEWL into the *AGA2* locus

The *AGA2* disrupted **a**-strain (CAY3; Table 5.11) containing both the pBS39 and pRS315-*lox* HEWL plasmids was treated with galactose using two different protocols: the liquid galactose method (Section 5.3.7) and the solid galactose method (Section 5.3.6). These protocols were followed exactly except the selective media used throughout was SC without histidine and uracil.

Colonies that could not grow on geneticin (G418) were selected from the original YPD plate and transferred to 10ml of YPD. These cultures were grown overnight at 30°C and genomic DNA was prepared for each (Section 2.2.21). Individual genomic DNA samples were screened using three primer sets in separate *Taq* PCRs (Section 2.2.7b). The primer set Rpl30A(seq)/LysB(HindIII) analysed the 5' end, HOP2(seq)/LysF(SphI) checked the 3' end and RPL30A(seq)/HOP2(seq) amplified the total *AGA2* locus (Table 5.10; Figure 5.16). The PCR conditions for the amplification of the 5' and 3' ends were: [94°C, 1min; 50°C, 1min; 72°C, 2min) x 30; 72°C 10min] and the parameters for the amplification of the total *AGA2* locus were

exactly the same except the elongation time was four minutes. Agarose gel electrophoresis (Section 2.2.8) was used to analyse the PCR products and one positive transformant was chosen for DNA sequencing. A 2.8 k.b.p. Expand™ amplified (Section 2.2.7c) PCR product was generated using the AGA2FUS(FOR)/HOP2(BACK) [Table 5.9] and the PCR conditions were: [94°C, 2min; (94°C, 15s; 52°C, 30s; 72°C, 2min) x 10; (94°C, 15s; 52°C, 30s; 72°C, 2min+20s/cycle) x 15; 72°C, 5min]. This fragment was purified using β-agarase (Section 2.2.9a) and the HEWL, *loxP* and *loxP 511* sites were sequenced with primers LysB(HindIII) and LysF(SphI). The **a**-strain contained the correctly integrated HEWL at its *AGA2* locus was named CAY4 (Table 5.11).

Figure 5.16 illustrates the primer sites used for the analysis and sequencing of HEWL at the 3' end of the *AGA2* locus.

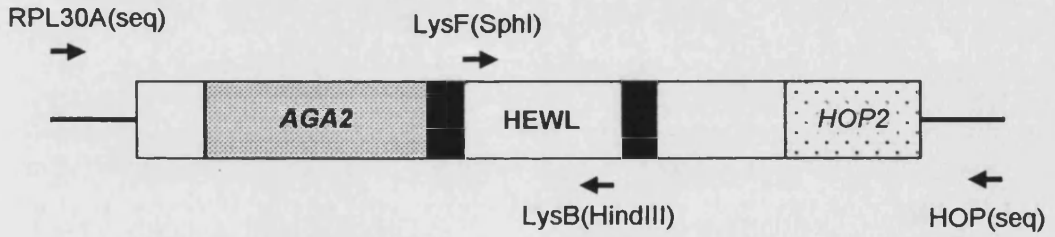


Table 5.11 Strains constructed in this chapter.

Strain Name	Genotype
CAY1	*MATα ura 3-52 leu2Δ1 sag1-Δ976::kanMX4*
CAY2	*MATα ura 3-52 leu2Δ1 sag1-Δ976*::scFv D1.3
CAY3	*MAT***a** *ura3-52 his3Δ200 aga2::kanMX4*
CAY4	*MAT***a** *ura3-52 his3Δ200 aga2*::HEWL

5.5 METHODS - CONSTRUCTION OF scFv D1.3 AND HEWL OVEREXPRESSION VECTORS

5.5.1 Construction of pYES2-*SAG1*-scFv D1.3 overexpression vector

The open reading frame of scFv D1.3 was amplified using pUC119His6mycXba-scFvD1.3 as a DNA template and the primer combination DF (Table 5.6) and DB(NotI) [Table 5.12]. An Expand™ PCR (Section 2.2.7c) was performed using the following cycling parameters: [94° C, 2min; (94°C, 15s; 50°C, 30s; 72°C, 50s) x 7; (94°C, 15s; 50°C, 30s; 72°C, 50s + 20s/cycle) x 8; 72°C, 5min]. The PCR buffers were removed (Section 2.2.9b) and the PCR product was examined by agarose gel electrophoresis (Section 2.2.8).

The vector constructed in Chapter 3, pYES2-*SAG1*-scFv B7 (Section 3.3.4) and the amplified scFv D1.3 were digested with *Hind* III (10units) and *Not* I (10units) overnight at 37°C (Section 2.2.12). The *Not* I/*Hind* III-digested pYES2-*SAG1*-scFv B7 vector was resolved on a low melting point agarose gel (Section 2.2.8) and the 6.7 k.b.p. fragment was purified (Section 2.2.9b). The restriction digestion buffers were also removed from the digested scFv D1.3 fragment (Section 2.2.9b). Both *Not* I/*Hind* III-digested products were quantified by agarose gel electrophoresis (Section 2.2.8) and ligated together (Section 2.2.13). After transformation (Section 2.2.15), ampicillin-resistant colonies were PCR screened (Section 2.2.7) using primers pYES2(FOR)/pYES2(BACK). Positive transformants were grown in 3ml of LB overnight (Section 2.2.1) and 'minipreps' were prepared (Section 2.2.17). A diagnostic restriction digest (Section 2.2.12) with *Hind* III (10units) and *Xho* I (20units) was performed on one plasmid DNA preparation to check the scFv D1.3 insert was present.

5.5.2 Transformation of the *SAG1* disrupted α-strain with pYES2-*SAG1*-scFv D1.3

The expression vector, pYES2-*SAG1*-scFv D1.3 (2 μg) was transformed into the α-strain containing the *loxP-kanMX4-loxP 511* cassette in its *SAG1* locus

(CAY1; Section 2.2.23). Ura^+ colonies were PCR screened using pYES(FOR)/pYES2(BACK) and positive transformants were grown in 1ml of SC lacking uracil [containing 2% (w/v) raffinose] for 24 hours (30°C). Glyercol stocks were made and a 100µl aliquot was plated onto a SC without uracil/ 2% (w/v) raffinose agar plate.

Table 5.12 - Sequences of the oligonucleotides used in the amplification of HEWL and scFv D1.3 in the construction of the overexpression vectors. Enzyme recognition sites are in bold.

Primer Name	Sequences (5'-3')	Description (5'-3')
DB (NotI)	CTGAGATACT**GCGGCCG** **C**CCGTTTGATCTCGAGCT T	10 n.t. overhang; a *Not* I site; 18 n.t. of homology to the 3' end of scFv D1.3 (non-coding strand).
LyspYES2(F)	CCCCCC**TGGCCA**GTGCT AAAGTCTTTGGACGATGT	6 n.t. overhang; a *Msc* I site; 5 n.t. encoding the end of the α-agglutinin leader sequence; 18 n.t. of homology to the 5' end of HEWL (coding strand).
LyspYES2(B)	CTGAGATACT**GCGGCCG** **C**CAGCCGGCAGCCTCT	10 n.t. overhang; a *Not* I site; 15 n.t. of homology to the 3' end of HEWL (non-coding strand).

5.5.3 Construction of pYES2-*SAG1*-HEWL overexpression vector

The HEWL open reading frame was amplified using pUC8-HEWL as a DNA template and the primers, lyspYES2(F)/lyspYES2(B) [Table 5.12]. An Expand™ PCR (Section 2.2.7c) was performed and the cycling parameters were: [94°C, 2min; (94°C, 15s; 50°C, 30s; 72°C, 45s) x 7; (94°C, 15s; 50°C, 30s; 72°C, 45 s + 20s/cycle) x 8; 72°C, 5min]. The amplified HEWL fragment and pYES2-*SAG1*-scFv B7 were digested with *Msc* I (3units) and *Not* I (10units) overnight at 37°C (Section 2.2.12). The *Msc* I/*Not* I-digested vector was resolved on a low melting point agarose gel (Section 2.2.8) and

recovered using a DNA extraction kit (Section 2.2.9b). The *Msc* I/*Not* I-digested HEWL was also purified (Section 2.2.9b) and both fragments were ligated together (Section 2.2.13). The ligation mixture was transformed into *E. coli* cells (Section 2.2.15) and the resulting colonies were PCR screened (Section 2.2.16) using pYES(FOR)/pYES2(BACK). Positive transformants were grown in 3ml of LB (Section 2.2.1) overnight and glycerol stocks were prepared (Section 2.2.3). Plasmid DNA was prepared from one overnight culture and transformed directly into *S. cerevisiae*.

5.5.4 Transformation of the *AGA2* disrupted strain with pYES2-*SAG1*-HEWL

The expression vector, pYES2-*SAG1*-HEWL (~2μg) was transformed into the **a**-strain containing the *loxP 511-kanMX4-loxP* cassette in its *AGA2* locus (CAY3; Section 2.2.23). Ura^+ colonies were PCR screened using pYES(FOR)/pYES2(BACK) and positive transformants were grown in 1ml of SC lacking uracil [containing 2% (w/v) raffinose] for 24 hours (30°C). Glyercol stocks were made and a 100μl aliquot was plated onto a SC without uracil/ 2% (w/v) raffinose agar plate.

5.6 RESULTS - CONSTRUCTION OF THE scFv D1.3-*SAG1* GENE FUSION

5.6.1 Construction of the *loxP-kanMX4-loxP 511 SAG1* disruption cassette

The construction of the *kanMX4* marker flanked by two regions of homology to the *SAG1* locus was constructed using a 'splicing by overlap extension' reaction (Figure 5.8).

In the first step, two DNA fragments encoding the promoter and the 3' anchorage domain of the *SAG1* locus were individually amplified from yeast genomic DNA. The sizes of both the generated PCR fragments were determined by agarose gel electrophoresis (Figure 5.17) and were in good agreement with the predicted theoretical sizes i.e. 1008 b.p. for the promoter region of *SAG1* locus and 1057 b.p. for the 3' half of the α-agglutinin gene. These DNA fragments were also designed to encode *loxP* and *loxP 511* sites which flanked either side of the *kanMX4* cassette.

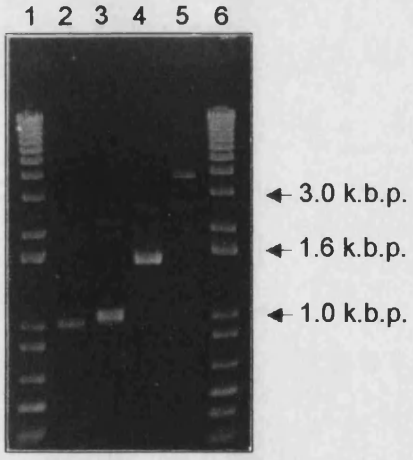


Figure 5.17 A 1.0% agarose gel showing the various fragments used in the construction of *loxP-kanMX4-loxP 511 SAG1* disruption cassette (Figure 5.8). Lanes 1 and 6, 1 k.b.p. DNA marker; Lane 2, *SAG1* promoter (1008 b.p.); Lane 3, C-terminal *SAG1* anchorage domain (1057 b.p.); Lane 4, *Not* I-digested *kanMX4* cassette (1.5 k.b.p.); Lane 5, *loxP-kanMX4-loxP 511 SAG1* disruption cassette (3.6 k.b.p.).

In the second PCR, these two DNA fragments were linked to the *Not* I digested-*kanMX4* cassette by a 'splicing by overlap extension' reaction (Horton *et al.*, 1989). This was possible due to the presence of the 25 nucleotides of homology each of the DNA fragments contained to the *kanMX4* cassette. A *SAG1* disruption cassette of 3.6 kilobases was obtained (Figure 5.17) which was in agreement with the expected size.

5.6.2 Integration of *SAG1* disruption cassette into the α-strain, FY10

The *loxP-kanMX4-loxP 511 SAG1* disruption cassette was successfully integrated into α-strain, FY10, at the 'N-terminal/C-terminal' junction of the *SAG1* locus, using a lithium acetate/polyethylene glycol transformation procedure (Section 2.2.24). A transformation efficiency of 63 geneticin-resistant ($G418^r$) transformants per μg of DNA was achieved.

PCR analysis of three $G418^r$ colonies showed that two of these had the *kanMX4* disruption cassette correctly integrated into the *SAG1* locus (Figure 5.18). As expected, the wild-type genomic DNA control did not produce any amplified products (data not shown).

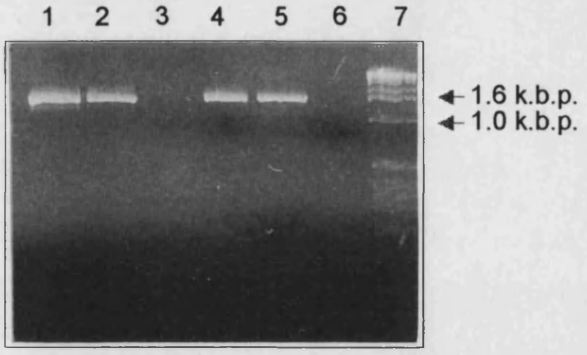


Figure 5.18 A 1.5% agarose gel confirming the integration of the *loxP-kanMX4-loxP 511* cassette into the N-terminal/C-terminal junction of the *SAG1* locus (Figure 5.10) for two $G418^r$ transformants. Lanes 1 and 2, Analysis at the 5' (promoter) end of the *SAG1* locus (1.4 k.b.p.); Lane 3, No DNA control; Lanes 4 and 5, PCRs confirming the integration at the 3' (C-terminal) end of the *SAG1* locus (1.5 k.b.p.); Lane 6, No DNA control, Lane 7, 1 k.b.p. DNA marker.

One of these transformants was examined by DNA sequencing and both the novel joints created by the homologous recombination event were correctly formed. Sequencing analysis also showed the *loxP* and *loxP 511* sites flanked either side of the kanamycin cassette and contained no DNA polymerase errors. The 3' half of the α-agglutinin gene was compared to the published sequence (Hauser and Tanner, 1989; SGD) and was found to be mutation-free. Additionally, the two extra nucleotides after the *loxP 511* site were present, therefore maintaining the desired reading frame for the scFv D1.3-*loxP 511*/α-agglutinin gene fusion (Figure 5.9).

5.6.3 Construction of the pYX243-*lox* D1.3 vector

The *GAL1* promoter was removed from the pYX243 vector using an inverse PCR (Triglia *et al.*, 1988). An 8 kilobase PCR product was obtained, which also incorporated a *loxP* site and a *Hind* III recognition site (Figure 5.19). This linear fragment was digested with *Hind* III/*Xho* I and a 56 b.p. *loxP 511* oligonucleotide cassette with *Hind* III/*Xho* I cohesive ends was inserted by ligation. The most efficient transformation was achieved using a molar ratio of 50:1 (insert:vector). Eight ampicillin-resistant colonies were PCR screened from the 50:1 transformation plate and a PCR product of approximately 370 b.p. in size was achieved for each. These PCR products were ethanol precipitated and digested with *Sph* I. The two expected DNA fragments of 146 b.p. and 223 b.p. were observed for five transformants, confirming that the two *loxP 511* oligonucleotides had annealed correctly and the *Sph* I recognition site had been created (data not shown).

The scFv D1.3 open reading frame (ORF) was amplified using primers DF and DB(SphI), and pUC119His6mycXba-scFvD1.3 as a DNA template. An 804 b.p. single PCR product was obtained, which was comparable with the expected size (Figure 5.19).

The scFv D1.3 fragment and the vector, pYX243-*loxP*/*loxP 511* were digested with *Hind* III/*Sph* I and then ligated together. Transformation of the construct produced several hundred plaques of which sixteen were PCR screened for the scFv D1.3 insert. Fourteen transformants were found to

be positive and DNA sequencing of this final pYX243-*lox* D1.3 construct exposed a problem with mutations in the scFv D1.3 gene. Of the three constructs sequenced only one vector was found to be mutation-free and the other two plasmids contained deleterious point mutations at various positions (Table 5.13). There was also an encoded amino acid in the amplified scFv D1.3 antibody fragment which was different ($Lys^{100} \rightarrow Glu$) from the published sequence (Kabat *et al.*, 1991). This difference was present in all the constructs sequenced and therefore was highly unlikely to be a polymerase error introduced during the amplification process. In fact, the supplied *E. coli* expression vector, pUC119His6mycXba-scFv D1.3 was analysed by DNA sequencing and was found to contain the $Lys^{100} \rightarrow Glu$ mutation. Sequencing of the final pYX243-*lox* D1.3 construct also determined the α-agglutinin signal sequence and the *loxP* and *loxP 511* sites were mutation-free. Additionally, the stop codon had been removed from the end of scFv D1.3 gene as intended. The extra adenine nucleotide after the *Sph* I recognition sequence was in place to maintain the in-frame fusion to the 3' half of the α-agglutinin gene (Figure 5.9).

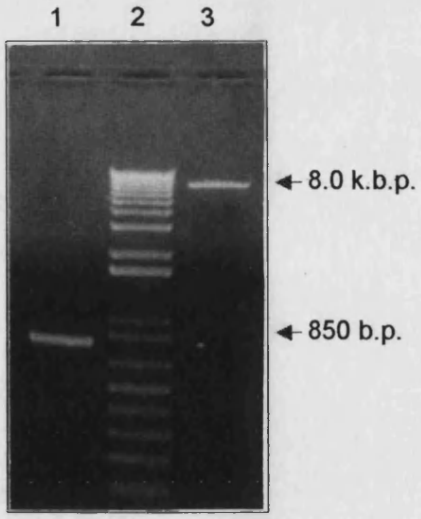


Figure 5. 19 A 0.8% agarose gel showing the results of the PCR amplification of the scFv D1.3 antibody fragment and vector, pYX243-*loxP*. Lane 1, scFv D1.3 (804 b.p.); Lane 2, 1 k.b.p. marker; Lane 3, pYX243-*loxP* (8 k.b.p.).

Table 5.13 Mutations found in the scFv D1.3 open reading frame. The numbering of the amino acid residues refers to the assembled scFv D1.3/α-agglutinin gene fusion (Figure 5.9).

Clone	Chain	Region	Residue
1	Heavy	FR3	Tyr^{113} → Asn
	Light	CDR1	Asn^{181} → Ser
2	Light	FR3	Gln^{229} → His

5.6.4 Verification of the Cre expression vector, pBS39

The Cre expression vector, pBS39 was digested with *Xho* I and gave the two diagnostic DNA fragments of 5.5 k.b.p. and 4.1 k.b.p. when analysed by agarose gel electrophoresis (data not shown).

The Cre recombinase gene sequence was also analysed and found to be mutation-free when compared to the published sequence (Sternberg *et al.*, 1986).

5.6.5 Cre-mediated integration of scFv D1.3 into the *SAG1* locus using a 2μ plasmid

Plasmids pBS39 and pYX243-*lox* D1.3 were co-transformed into the α-strain containing the *loxP-kanMX4-loxP 511* cassette in its *SAG1* locus (CAY1; Table 5.11). Approximately 300 Ura^+ Leu^+ transformants were obtained and three colonies were PCR screened to confirm both the plasmids were present.

The expression plasmid pBS39 containing the Cre recombinase gene under the control of the *GAL1* promoter was induced using the liquid galactose induction protocol. Cells were grown to mid-log phase in synthetic complete (SC) media containing raffinose as the carbon source, and then induced for 24 hours with galactose. After this time, Cre recombinase expression was stopped by transferring the cells to non-selective media containing glucose (YPD). Theoretically, the pYX243-*loxP/loxP 511* plasmid should now have the *kanMX4* cassette (Figure 5.3; Stage 4). This vector contains elements that allow autonomous replication

in yeast, therefore could make the recombinants geneticin-resistant ($G418^r$). Consequently, to detect colonies with a geneticin-sensitive ($G418^s$) phenotype, this plasmid was removed by 3-4 days growth in non-selective medium.

Removal of the *kanMX4* cassette from the *SAG1* locus was detected by transferring the cells which had been grown in non-selective media for 3-4 days onto YPD plates and replica plating these colonies onto YPD/G418 plates.

The percentage of recombination between the *loxP* and *loxP 511* sites was determined from the ratio of the number of colonies which grew on the YPD/G418 plates to the number of colonies on the YPD plates (Table 5.14).

Approximately 17% of galactose-induced cells containing plasmids pBS39 and pYX243-*lox* D1.3 had lost the *kanMX4* cassette. This loss appeared to be specific to Cre recombinase expression because in glucose media, where little Cre recombinase activity was expected, the *kanMX4* cassette was not removed. From this limited set of data, the expression level of Cre recombinase seemed to fluctuate in different induction experiments. This data also indicates that undesirable recombination was occurring between *loxP* and *loxP 511* sites. Hypothetically, the disrupted *SAG1* α-strain (CAY1) containing only pBS39 should not generate $G418^s$ colonies when grown in media containing galactose as a carbon source.

Fourteen $G418^s$ Leu^- colonies were randomly picked for further PCR analysis from the double plasmid (pBS39 and pYX243-*lox* D1.3) galactose induction experiments. Expected DNA fragments of 2.0 k.b.p. for the 5' promoter end and 1.4 k.b.p. for the 3' end of the *SAG1* locus were obtained (Figure 5.20) for six of these colonies. This indicated the scFv D1.3 gene had been correctly integrated into the *SAG1* locus using Cre recombinase. One of these positive colonies was analysed by DNA sequencing and this showed that the scFv D1.3 gene was correctly placed between the promoter and 3' half of the α-agglutinin gene in the yeast genome (Figure 5.9). The DNA encoding the α-agglutinin leader sequence, scFv D1.3

antibody fragment and the C-terminal half of the α-agglutinin was mutation-free.

Table 5.14 Table showing the percentage of observed[a] recombination (% Rec.) between *loxP/loxP 511* sites on the 2μ plasmid and *SAG1* locus in a number of Cre expression experiments (30°C).

Plasmids [b]	Carbon Source	Induction method/ time	No. of G418^S colonies/Total no. of colonies	% Rec.
pBS39/pYX243-*lox* D1.3	Galactose	Liquid/ 24hrs	44/463	10
"	"	"	75/330	23
pBS39 only	Galactose	Liquid/ 24 hrs	31/400	8
pBS39/pYX243-*lox* D1.3	Glucose	Liquid/ 24 hrs	0/300	0

a - The recorded percentage recombination does not include any double recombination events i.e. the *kanMX4* cassette can be replaced by the scFv D1.3 at the *SAG1* locus and a further recombination event can occur where the genomic scFv D1.3 is substituted for the *kanMX4* cassette from the exchange vector. Any double (or quadruple etc.) recombination events that occur in these experiments are undetectable.

b - All plasmids were transformed into strain CAY1 (*MATα ura3-52 leu2Δ1 sag1-Δ976::kanMX4*).

PCR amplification of the entire *SAG1* locus of the remaining eight G418^S Leu$^-$ colonies indicated that they had lost the *kanMX4* cassette without acquiring the scFv D1.3 gene since PCR products of 2.5 k.b.p. instead of the expected size of 3.3 k.b.p were observed. These results indicate that the *loxP* and *loxP 511* sites were undesirably recombining and

the *kanMX4* cassette could be lost from the *SAG1* locus without exchange for the scFv D1.3 gene.

Two colonies were also PCR screened from the galactose induction experiment containing the Cre recombinase expression vector (pBS39) only. These colonies were also shown to have lost the *kanMX4* cassette. Analysis of the DNA sequence of one of these colonies exposed a single *loxP 511* site flanked by the promoter and the 3' half of the α-agglutinin gene. This result confirmed the speculation that undesirable recombination was occurring between the *loxP* and *loxP 511* sites.

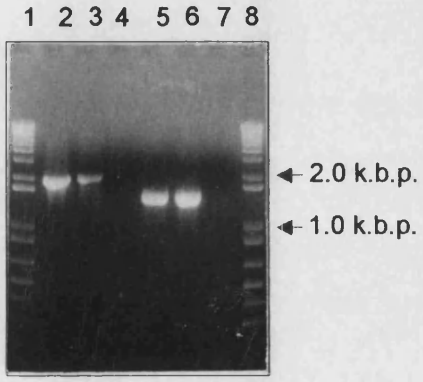
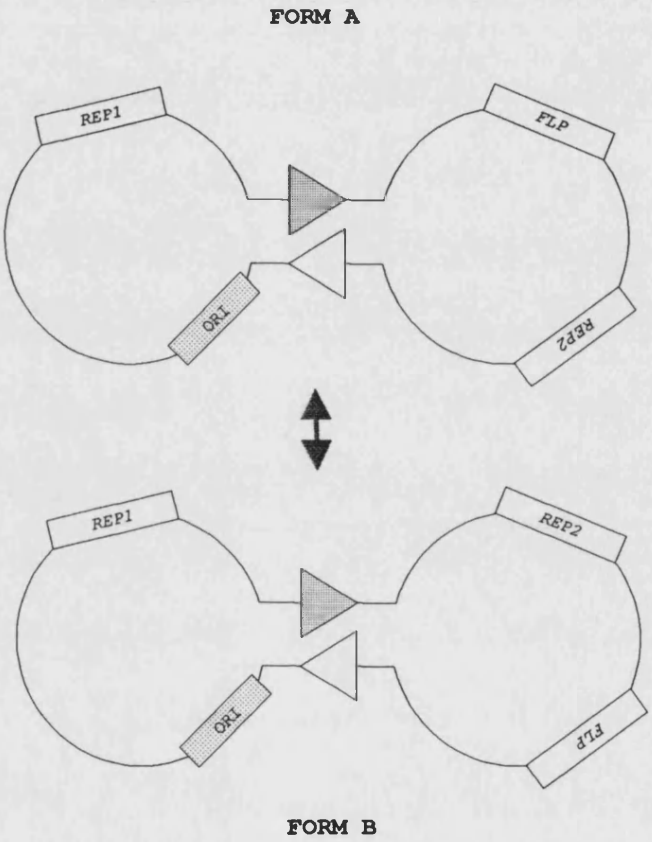


Figure 5.20 A 0.8% agarose gel verifying the Cre-mediated integration of scFv D1.3 gene into the *SAG1* locus of strain CAY1. The PCR analysis was performed on two individual colonies at the 5' and 3' ends of the *SAG1* locus (Figure 5.12). Lanes 1 and 8, 1 k.b.p. DNA marker; Lanes 2 and 3, 5' (promoter) end of the *SAG1* locus (2. k.b.p); Lane 4, No DNA control; Lanes 5 and 6, 3' (C-terminal) end of the *SAG1* locus (1.4 k.b.p.); Lane 7, No DNA control.

Due to the *loxP* and *loxP 511* sites recombining at the *SAG1* locus (or on the exchange vector), a centromeric vector was constructed. It was reasoned that the conversion between the A and B isomeric forms of the native 2μ plasmid (Figure 5.21) and the exchange vector, pYX243-*lox* D1.3, maybe producing intermediates which were affecting the fidelity of the Cre-*lox* system. Yeast shuttle vectors containing inverted repeats have

been shown to exist as a variety of recombinants with the native 2μ plasmid (Gerbaud *et al.*, 1979).

Figure 5.21 The A and B forms of the 2μ plasmid. The inverted repeats are shown as triangles.



The native 2μ plasmid (6.3 k.b.p.) is present in most *S. cerevisiae* strains at about 50 copies per haploid genome (Broach, 1982). It can interconvert between two distinct states by an autocatalysed, intramolecular recombination. Recombination occurs between the two inverted repeats (599 b.p.) and is catalysed by the protein encoded by the 2μ circle FLP gene. Additionally, the 2μ circle contains an origin of replication (ORI) for autonomous replication and genes, *REP1* and *REP2*, which are essential for stable propagation and high copy number.

5.6.6 Construction of the pRS315-*lox* D1.3 vector

The *Kluyvermyces lactis URA3* gene was removed from plasmid TG5756 (Langle-Rouault *et al.*, 1995) by performing a *Not* I/*Apa* I digest. This generated the digested centromeric vector, pRS315-*Not* I/*Apa* I.

The *loxP*-scFv D1.3-*loxP 511* DNA fragment was amplified using primers, pRS315(ApaI) and pRS315(NotI), and the construct pYX243-*lox* D1.3 as a DNA template. A single PCR product of approximately 900 b.p. was obtained. This PCR product was digested with *Not* I/*Apa* I and cloned into pRS315-*Not*I/*Apa* I. Transformation of the construct produced several hundred plaques of which ten were PCR screened for the scFv D1.3 gene. Nine were found to be positive and DNA sequencing analysis of one of these transformants revealed a deleterious point mutation [Ser^{222} → Pro (numbering corresponds to scFv D1.3-*SAG1* gene fusion; Figure 5.9)] in the open reading frame of the scFv D1.3 gene. The *loxP* and *loxP 511* sites were mutation-free, therefore the plasmid was used. This centromeric vector was only required for a parallel study in the percentage of recombination between *loxP*/*loxP 511* sites compared to the 2μ plasmid, therefore the mutation was unimportant.

5.6.7 Cre-mediated integration of scFvD1.3 into the *SAG1* locus using the centromeric plasmid

Plasmids pBS39 and pRS315-*lox* D1.3 were co-transformed into the α-strain containing the *loxP-kanMX4-loxP 511* cassette in its *SAG1* locus (CAY1; Table 5.11). Approximately 200 Ura^+ Leu^+ transformants were obtained and six colonies were PCR screened to confirm both plasmids were present.

The Cre recombinase expression vector was induced using two different protocols: the solid and liquid galactose induction methods. The liquid galactose induction method (Section 5.3.7) was analogous to the protocol used for the 2μ plasmid (Section 5.6.5).

Cre recombinase expression was also induced on minimal media supplemented with either galactose or glucose (Solid galactose induction method; Section 5.3.6). These plates were incubated for 4 days and then

individual colonies were picked from each plate and resuspended in non-selective glucose media (YPD) to suppress Cre recombinase expression. These cultures were grown for 3-4 days in non-selective media to remove the pRS315-*lox-kanMX4* plasmid (Figure 5.4; discussed in Section 5.6.5).

Removal of the *kanMX4* cassette from the *SAG1* locus was detected by transferring cells which had grown in non-selective media for 3-4 days onto YPD plates and replica-plating these colonies onto YPD/G418 plates.

The solid galactose induction experiments with plasmids, pBS39 and pRS315-*lox* D1.3 resulted in approximately 27% of the colonies becoming geneticin-sensitive (Table 5.15). PCR analysis of twelve $G418^s$ Leu^- colonies randomly picked from these plates indicated that none of the colonies contained the scFv D1.3 gene in their *SAG1* loci. Further examination showed that all twelve colonies had removed the *kanMX4* cassette from the *SAG1* locus without acquiring the scFv D1.3 gene. The glucose control plate did not generate any $G418^s$ colonies. Therefore, the loss of the *kanMX4* cassette appears to be related to Cre recombinase expression.

The liquid galactose induction experiments with plasmids pBS39 and pRS315-*lox* D1.3 showed that the number of $G418^s$ colonies increased with incubation time (Table 5.15). Ten $G418^s$ Leu^- colonies from each time point were further analysed by the PCR and only one colony from the 5-hour time point had incorporated the scFv D1.3 gene into its *SAG1* locus. At the 24-hour time point, three colonies had correctly integrated the scFv D1.3 gene. Consequently, from this limited set of data, it can be deduced that induction time appears to have no effect on the fidelity of the Cre-*lox* system. Further experiments are required to confirm this conclusion.

The results obtained using the centromeric vector supports the data acquired from the 2μ plasmid (Section 5.6.5), namely that the *loxP* and *loxP 511* sites appear to be illegitimately recombining at the *SAG1* locus.

Table 5.15 Table showing the percentage of observed[a] recombination (% rec.) between *loxP/loxP 511* sites on the CEN plasmid and the *SAG1* locus in a number of Cre induction experiments (30°C).

Plasmid(s)[b]	Carbon Source	Induction method/ time	No. of G418[s] colonies/ total no. of colonies	% rec.
pBS39/pRS315-*lox* D1.3	Galactose	Solid/ 4 days	60/203	30
"	Galactose	Solid/ 4 days	63/313	20
"	Galactose	Solid/ 4 days	62/210	30
"	Glucose	Solid/ 4 days	0/250	0
pBS39/pRS315-*lox* D1.3	Galactose	Liquid/ 5 hours	81/790	10
"	Galactose	Liquid/ 24 hours	170/555	31

a - The recorded percentage recombination does not include any double recombination events i.e. the *kanMX4* cassette can be replaced by the scFv D1.3 at the *SAG1* locus and a further recombination event can occur where the genomic scFv D1.3 is substituted for the *kanMX4* cassette from the exchange vector. Any double (or quadruple etc.) recombination events that occur in these experiments are undetectable.

b - All plasmids were transformed into strain CAY1 (*MATα ura3-52 leu2Δ1 sag1-Δ 976::kanMX4*).

5.7 RESULTS - CONSTRUCTION OF THE *AGA2*-HEWL GENE FUSION

5.7.1 Construction of *loxP 511-kanMX4-loxP AGA2* disruption cassette

The *kanMX4* cassette flanked by two regions of homology to the 3' end of the *AGA2* locus was constructed by a 'splicing by overlap extension' reaction.

Firstly, two DNA fragments encoding the *AGA2* or *HOP2* ORFs and surrounding sequences were individually amplified from yeast genomic DNA. The sizes of the amplified DNA fragments (Figure 5.22) were in good agreement with the expected sizes i.e. 885 b.p. for the *AGA2* DNA fragment and 1343 b.p. for the *HOP2* DNA fragment.

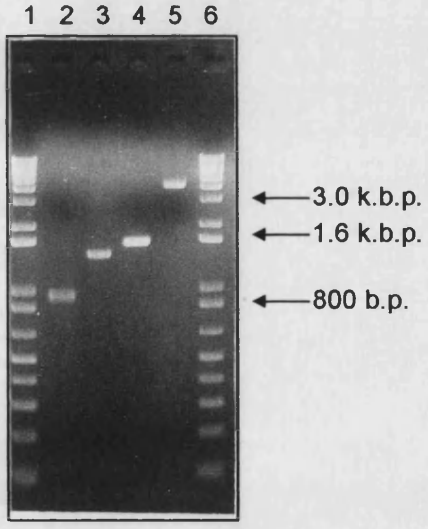


Figure 5.22 A 0.8% agarose gel showing the various fragments used in the construction of the *loxP 511-kanMX4-loxP AGA2* disruption cassette (Figure 5.13). Lanes 1 and 6, 1 k.b.p. marker; Lane 2, Amplified *AGA2* fragment (885 b.p.); Lane 3, *HOP2* fragment (1343 b.p.); Lane 4, *Not* I-digested *kanMX4* cassette (1.5 k.b.p.); Lane 5, Assembled *loxP 511-kanMX4-loxP AGA2* disruption cassette (3.7 k.b.p.).

In the second PCR, these two fragments were annealed to the *Not* I-digested *kanMX4* cassette by a 'splicing by overlap extension' reaction and a *loxP 511-kanMX4-loxP 511 AGA2* disruption cassette of 3.7 kilobase pairs was obtained (Figure 5.22).

5.7.2 Integration of the *AGA2* disruption cassette into the a-strain, FY22

The *loxP 511-kanMX4-loxP AGA2* disruption cassette was successfully transformed into the **a**-strain (FY22). A transformation efficiency of 96 geneticin-resistant ($G418^r$) transformants per μg of DNA was achieved using a lithium acetate/polyethylene glycol transformation procedure (Section 2.2.24).

PCR analysis of two of the $G418^r$ colonies showed that both contained the *kanMX4* disruption cassette integrated into the *AGA2* locus (Figure 5.23).

Figure 5.23 A 0.8% agarose gel confirming the integration of the *loxP 511-kanMX4-loxP* at the 3' terminus of the *AGA2* ORF. The PCR analysis was performed on two individual $G418^r$ colonies at the 5' and 3' integration sites (Figure 5.15). Lanes 1 and 6, 1 k.b.p. marker; Lanes 2 and 3, 5' integration site of the *AGA2* locus (1178 b.p.); Lane 4, No DNA control; Lanes 7 and 8, 3' integration site of the *AGA2* locus (1644 b.p.); Lane 9, No DNA control.

Analysis of the DNA sequence of one of the transformants showed the *loxP 511* and *loxP* sites flanked the *kanMX4* cassette and the total disruption cassette was correctly integrated at the 3' terminus of the *AGA2* locus. The stop codon had been deliberately removed from the 3' end of the *AGA2* ORF in the amplification process, therefore maintaining the correct reading frame for the *AGA2-loxP 511*-HEWL gene fusion (Figure 5.14). Both the *loxP/loxP 511* sites and the *AGA2* coding sequence were mutation-free.

5.7.3 Construction of the pRS313-*lox* HEWL vector

The hen egg white lysozyme (HEWL) open reading frame was amplified using the primer pair, lysF(SphI) and lysB(HindIII), and the vector, pUC8-HEWL as a DNA template. A single PCR fragment of approximately 400 b.p. in size was produced (Figure 5.24) and digested with *Sph* I and *Hind* III, along with the pYX243-*loxP/loxP 511* vector. After ligation, transformation of the construct produced several hundred plaques, of which seven were screened for the HEWL insert. Five transformants were found to be positive and one of these was analysed by DNA sequencing. The lysozyme and *loxP/loxP 511* sites were found to be mutation-free and the designed extra stop codon at the 3' end of the lysozyme gene was present. The extra adenine nucleotide at the 5' end of the lysozyme gene was also in place to maintain the correct frame for the *AGA2*/HEWL gene fusion (Figure 5.14). This plasmid was only used for the construction of the pRS313-*lox* HEWL vector.

The lysozyme flanked by the two *loxP/loxP 511* sites was amplified from the previous construct using the primer combination pRS315(ApaI) and pRS315(NotI). A single 503 b.p. PCR product was obtained and digested with the enzymes *Apa* I and *Not* I, along with the centromeric vector, pRS313. These digested fragments were ligated together and seven of the resulting ampicillin-resistant colonies were screened for insert. Six transformants were positive and DNA sequencing confirmed the *loxP/loxP 511* sites flanked the lysozyme gene. No DNA polymerase errors were discovered in the desired construct.

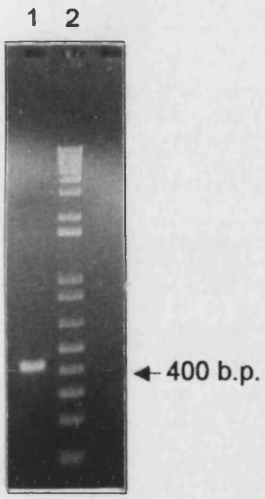


Figure 5.24 A 1.5% agarose gel showing the PCR amplification of hen egg white lysozyme. Lane 1, HEWL product (421 b.p.); Lane 2, 1 k.b.p. DNA Marker.

5.7.4 Cre-mediated integration of HEWL into the *AGA2* locus

The vectors pRS313-*lox* HEWL and pBS39 were co-transformed into the **a**-strain containing the *loxP 511-kanMX4-loxP* cassette at the 3' terminus of its *AGA2* locus (CAY3; Table 5.11). Approximately two hundred Ura^+ His^+ transformants were obtained and all colonies screened contained both plasmids.

The Cre recombinase expression vector, pBS39, was induced using both the liquid and solid galactose induction protocols (Sections 5.3.6 and 5.3.7). The removal of the *kanMX4* cassette from the *AGA2* locus was detected as previously described in Section 5.6.5 and Table 5.16 summarises the recombination results obtained.

When disrupted **a**-cells (CAY3) containing plasmids pBS39 and pRS313-*lox* HEWL were induced on solid media, approximately 49% of the colonies became sensitive to geneticin. Analysis of twelve $G418^s$ His^- colonies by the PCR showed that all transformants screened had lost the *kanMX4* cassette without acquiring the lysozyme gene.

Liquid induction data also demonstrated undesirable recombination between the *loxP* and *loxP 511* sites at the *AGA2* locus. The number of colonies that were $G418^s$ increased over the course of the induction

experiment and was specific to galactose induction. The number of $G418^S$ colonies obtained by the solid galactose induction method was three-fold higher than the number of $G418^S$ colonies achieved for the liquid galactose induction experiment. This result can also be assumed to be a product of a longer galactose induction time. Twenty-two $G418^S$ His^- colonies from the 24-hour time point and twelve $G418^S$ his^- colonies from the 5-hour time point were further analysed by the PCR. Again, most of the colonies had lost the *kanMX4* cassette between the recombination sites at the *AGA2* locus without gaining the HEWL gene. Only two colonies from the 24h-time point had the lysozyme gene correctly integrated into the *AGA2* locus (Figure 5.25).

DNA sequencing of one of these colonies proved that the lysozyme gene was correctly exchanged for the *kanMX4* cassette at the 3' terminus of the *AGA2* locus, inbetween the *loxP 511* and *loxP* sites. The *loxP 511* recombination site formed a peptide linker between the coding region of *AGA2* and HEWL. The desired 3' *AGA2* gene fusion with lysozyme was in the correct reading frame and mutation-free (Figure 5.14).

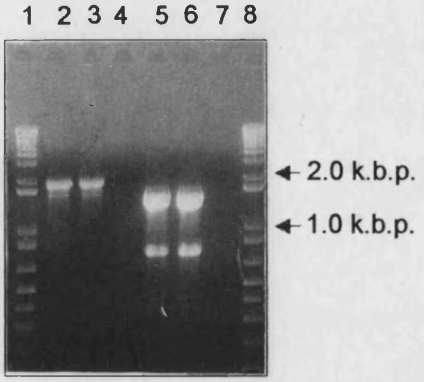


Figure 5.25 A 0.8% agarose gel verifying the Cre-mediated integration of HEWL gene into the 3' end of the *AGA2* locus of strain, CAY3. PCR analysis was performed on two individual $G418^S$ colonies to verify exchange had taken place at the 5' and 3' *AGA2* recombination sites (Figure 5.16). Lanes 1 and 8, 1 k.b.p. DNA Marker; Lanes 2 and 3, 3' recombination end of the *AGA2* locus (2.0 k.b.p.); Lane 4, No DNA control, Lanes 5 and 6, 5' recombination end of the *AGA2* locus (1.5 k.b.p.); Lane 7, No DNA control.

Table 5.16 Table showing the percentage of observed recombination[a] (% rec.) between *loxP/loxP 511* sites on the centromeric plasmid and the *AGA2* locus in a number of Cre induction experiments (30°C).

Plasmids [b]	Carbon Source	Induction method/ time	No. of G418[R] colonies/Total no. of colonies	% rec.
pBS39/pRS313-*lox* HEWL	Galactose	Solid/ 4 days	64/125	51
"	"	"	136/293	46
"	"	"	268/529	51
"	Glucose	Solid/ 4 days	0/250	0
pBS39/pRS313-*lox* HEWL	Galactose	Liquid/ 5 hrs	67/650	10
pBS39/pRS313-*lox* HEWL	Galactose	Liquid/ 24 hrs	46/324	14
"	"	"	57/374	15
"	"	"	104/518	20
pBS39/pRS313-*lox* HEWL	Glucose	Liquid/ 24 hrs	0/300	0
pBS39 only	Galactose	Liquid/ 24 hrs	51/423	12
"	Glucose	Liquid/ 24 hrs	0/300	0

a - The recorded percentage recombination does not include any double recombination events i.e. the *kanMX4* cassette can be replaced by the HEWL gene at the *AGA2* locus and a further recombination event can occur where the genomic HEWL is substituted for the *kanMX4* cassette from the exchange vector. Any double (or quadruple etc.) recombination events that occur in these experiments are undetectable.

b - All plasmids were transformed into *S. cerevisiae* strain, CAY3 (*MAT***a** *ura3-52 his3Δ200 aga2::kanMX4*).

5.8 RESULTS - CONSTRUCTION OF scFv D1.3 AND HEWL OVEREXPRESSION VECTORS

5.8.1 Construction of pYES2-*SAG1*-scFv D1.3 expression vector and transformation into the *SAG1* disrupted α-strain

The scFv D1.3 antibody fragment was amplified from vector pUC119His6mycXba-scFv D1.3 using the primers DF and DB(NotI). A PCR product of approximately 800 b.p. in size was obtained and digested with *Hind* III/*Not* I, along with the plasmid pYES2-*SAG1*-scFV B7 (Section 3.4.4). After ligation, screening of seven ampicillin-resistant colonies revealed each one contained the scFv D1.3 insert. A diagnostic restriction digest with *Hind* III and *Xho* I produced the expected DNA fragment of approximately 750 b.p. in size. This confirmed the pYES2-*SAG1*-scFv D1.3 overexpression vector was constructed.

The expression vector, pYES2-*SAG1*-scFv D1.3 was transformed into the *SAG1* disrupted α-strain, CAY1. Numerous Ura^+ transformants were obtained and several colonies contained the vector when screened.

5.8.2 Construction of pYES2-*SAG1*-HEWL expression vector and transformation into the *AGA2* disrupted a-strain.

The hen egg white lysozyme gene was amplified from vector pUC8-HEWL with the primers, lyspYES2(F) and lyspYES2(B). A PCR product of approximately 400 b.p. in size was achieved and digested with *Msc* I/*Not* I. The vector pYES2-*SAG1*-scFv B7 (Section 3.4.4) was also digested with *Msc* I/*Not* I and these two fragments were ligated together. After transformation, several hundred ampicillin-resistant colonies were obtained and all colonies screened were positive for the lysozyme insert. The pYES2-*SAG1*-HEWL overexpression vector was successfully transformed into the *AGA2* disrupted a-strain, CAY3.

5.9 DISCUSSION

5.9.1 Integration of the agglutinin disruption cassettes

The PCR-generated disruption cassettes were a fast and effective way of producing mutants at the appropriate agglutinin loci. The efficiency of integration of the disruption cassettes was approximately 60-100 transformants per μg of DNA (Sections 5.6.2 and 5.7.2) which is comparable to published results (Wach, 1996). Analysis by PCR screening and DNA sequencing confirmed the dominant resistance selectable marker, *kanMX4*, flanked by the *loxP/loxP 511* sites, was at the correct location in both the agglutinin loci. Extra nucleotides were also present at the required positions in the integrated constructs to maintain the reading frames of the agglutinin gene fusions (Figures 5.9 and 5.14).

5.9.2 Construction of the exchange vectors

The construction of the exchange vectors containing either the HEWL or scFv D1.3 genes flanked by the *loxP/loxP 511* sites was also successful. Careful design of oligonucleotides containing restriction endonuclease recognition sites that were compatible with the chosen vector's unique recognition sites facilitated the cloning of the antigen and antibody DNA.

The sequencing of scFv D1.3 gene exposed a mutational problem although eventually a construct that was mutation-free was identified (Section 5.6.3). The constructs containing deleterious point mutations could have been used in a pilot experiment to demonstrate the effectiveness of the yeast 'germinal centre' strategy (Part II). A mixed population of α-cells displaying the native or either of the two mutated scFv D1.3 antibody fragments could have been created. Hopefully, after mating of this mixed α-cell population with the engineered **a**-cells displaying lysozyme, results would have indicated that the yeast α-cells expressing the native scFv D1.3 antibody fragments were able to mate with greater efficiency than α-cells displaying the mutated antibody fragments. Hence, this pilot experiment would have illustrated the ability of the yeast 'germinal

centre' reaction to distinguish between antibody fragments with different affinity constants for the chosen antigen.

Sequence analysis of the scFv D1.3 antibody fragment also discovered an amino acid residue ($Lys^{100} \rightarrow Glu$) which differed from the published sequence in all the pYX243-*lox* D1.3 constructs. Based on computer graphic models of the lysozyme-scFv D1.3 complex, this residue was considered to be unimportant in the antibody-antigen interaction due to the fact it was remote from the antibody recognition site (personal communication, Dr. S. Searle). Functional expression of the scFv D1.3 antibody fragment in *E. coli* suggests this assumption is correct (Section 6.4.1).

5.9.3 Cre-mediated integration

Cre-mediated integration of the antigen and antibody DNA into the relevant agglutinin loci was achieved, although problems with the Cre-*lox* recombination system were encountered. All results indicated that the *loxP* and *loxP 511* sites were illegitimately recombining at the agglutinin loci and the *kanMX4* cassettes were lost without the incorporation of the scFv D1.3/HEWL DNA. It is also possible that the *loxP* and *loxP 511* sites were additionally recombining on the vectors and removing the HEWL/scFv D1.3 genes before exchange could ever take place. Preliminary results investigating the ability of the exchange vector to remove the HEWL/scFV D1.3 DNA indicates that this event can take place (personal communication, Bryn Davies, this laboratory). Using a centromeric vector instead of a 2μ plasmid did not overcome the undesirable *loxP* x *loxP 511* recombination problems. Hence, it can be assumed that any intermediates, which could occur between the native 2μ plasmid and the exchange vector (Section 5.6.5; Figure 5.21), were not affecting the fidelity of the Cre-*lox* system.

Recombination appeared to be a consequence of Cre recombinase expression due to the fact that manipulated strains did not generate any geneticin-sensitive colonies in experiments where Cre recombinase expression was inhibited by growth on medium containing glucose. The

percentage recombination between *loxP* (or *loxP 511*) sites increased with induction time, presumably because of the increase in the expression level of Cre recombinase. The maximum percentage recombination (~50%) was observed in experiments using solid media after a four-day induction period (Section 5.7.4). All the colonies analysed from these experiments had removed the *kanMX4* cassette without acquiring the desired DNA. It was therefore speculated that a longer induction time could be affecting the fidelity of the Cre recombinase system. Further examination of induction times using the liquid media method contradicted this view.

The maximum level of recombination achieved using the liquid media method was approximately 30%, although this level fluctuated greatly from experiment to experiment (Section 5.6.7). It was presumed that this fluctuation was due to the variation in the copy number of the Cre recombinase expression vector. The limited set of data available from the liquid media experiments demonstrated that the induction time had no affect on the fidelity of the Cre-*lox* system. Colonies analysed at the five-hour time point had not exchanged the *kanMX4* cassette more efficiently than colonies at the twenty-four hour time point.

The levels of recombination in all these experiments were much lower than published results where selectable markers were flanked by two homologous *loxP* sites (Sauer, 1987; Güldener *et al.*, 1996). The result of these studies reported approximately 90% of the relevant selectable marker gene had been removed from cells after an induction time of one hour. This may be an indication that the non-homologous *loxP* and *loxP511* are not recombining at the same efficiency in the genome as two homologous *loxP* sites.

Interestingly, recombination between non-homologous *loxP* and *loxR* sites in the yeast genome also results in a 90% loss of the selectable marker (Section 5.1.1; Ross-Macdonald *et al.*, 1997). Studies in bacteriophage λ indicate that *loxR* and *loxP* sites recombine at a rate 20- to 100-fold lower than two homologous *loxP* sites (Sternberg *et al.*, 1981). The fact that the non-homologous *loxP* and *loxR* sites recombine at the same efficiency as two homologous *loxP* sites in the yeast genome may

indicate that some degree of selectivity of the Cre recombinase enzyme is lost in a eukaryotic nucleus. This data appears to support the decrease in Cre recombinase fidelity observed in this project.

Intermolecular recombination of *loxP* sites has been reported to be less effective than intramolecular recombination (Abremski *et al.*, 1983). Hence, if the *loxP* and *loxP 511* sites are recombining undesirably, the recombination results obtained in this Chapter are not surprising i.e. intramolecular recombination is favoured, therefore the majority of colonies have lost the *kanMX4* cassette.

Cre-mediated integration of HEWL gene into the *AGA2* locus appears to be less effective than integrating scFv D1.3 DNA into the *SAG1* locus. There could be many reasons to explain this observation. For instance, the smaller size of the lysozyme gene could facilitate more intramolecular recombination because the *loxP/loxP 511* sites are physically closer, or possibly Cre recombinase activity is favoured at one of the chromosomal locations.

The factors affecting the fidelity of the Cre-*lox* recombination system could be numerous and were not investigated further in this project. The initial study, which determined that homology was required between *loxP* sites for efficient recombination, was performed in *E. coli* at 37°C (Hoess *et al.*, 1986). Consequently, the fact that these experiments were carried out in *S. cerevisiae* (an eukaryotic nucleus) at 30°C (not 37°C) maybe important.

5.9.4 Conclusions

Overall, the aims of this Chapter were achieved, but unexpected problems with the novel yeast Cre-*lox* recombination system were encountered. The maximum integration efficiency obtained using the Cre-*lox* system was appropriately 10^4 transformants per μg of plasmid DNA. The theoretical size of an antibody library using the Cre-*lox* system was predicted to be 10^6 transformants per μg of plasmid DNA (Section 5.1; Schiestl *et al.*, 1993). Although this was not achieved, the efficiency obtained in this Chapter is still two magnitudes higher than attainable by standard yeast integration

methods (1-100 transformants per μg of DNA; Wach, 1996). Further improvements to the Cre-*lox* system may generate the potential to create a small antibody library (~10^6 transformants) where moderate affinity antibodies (μM) could be isolated directly.

Both the antibody and antigen surface display systems, required for the yeast 'germinal centre' strategy, were produced using the Cre-*lox* system. The antibody DNA (scFv D1.3) successfully replaced the 5' α-agglutinin binding domain and an in-frame gene fusion with the 3' α-agglutinin anchorage domain was created (the antibody display system). The antigen display model was also constructed by fusing the HEWL DNA to the 3' terminus of the **a**-agglutinin binding domain gene. Both these engineered strains were individually analysed for the appropriate protein expression in the following Chapter and the artificial yeast 'germinal centre' reaction was also attempted.

Additionally, vectors were constructed for overexpression of scFv D1.3 and HEWL on the surfaces of α- and **a**-cells, respectively. Yeast strains transformed with these vectors were used as positive controls in the expression studies carried out in Chapter 6.

CHAPTER 6

ATTEMPTED MATING OF AGGLUTININ MUTANTS

6.1 INTRODUCTION

The overall aim of this part of the project is to produce an antigen-selection process which can distinguish between yeast cells displaying antibodies with different affinities for a chosen antigen on their cell surfaces (Part II). Antibody improvement in the human body occurs in the germinal centre via stepwise rounds of mutation and antigen-selection (Section 1.3.3). Our proposed strategy intends to mimic the germinal centre reaction as closely as possible by mating engineered α- and **a**-cells displaying antibody fragments or antigen on their cell surfaces, respectively (Figure 1.11).

The preceding chapter described the construction of both the *sag1-Δ976*::scFv D1.3 and *aga2*::HEWL genomic fusions in the α- and **a**-strains, respectively. This chapter details the individual expression of Aga2p-lysozyme and Sag1p-scFvD1.3 protein fusions on the surface of these engineered cells in both the presence and absence of pheromone. Exposure of cells to pheromone secreted by the opposite mating type (**a**-/α-factor) results in an increase in the number of agglutinin molecules on the cell surface [(2-5 x 10^4 agglutinin molecules per cell) Terrance and Lipke, 1987; Watzele *et al.*, 1988; Wojciechowicz and Lipke, 1989]. Additionally, the functionality of the scFv D1.3 and lysozyme on the surface of these engineered cells was examined using either biotinylated lysozyme or *Micrococcus lysodeikitus* cells, respectively.

Imitation of the germinal centre reaction was attempted by mating these **a**- and α-engineered cells (strains; CAY4 and CAY2) under conditions that did not promote stable cell-to-cell contact (liquid media; Lipke *et al.*, 1989) and then selecting for diploids. Theoretically, α-cells expressing the scFv D1.3 on their surface should form a stable cell-to-cell contact with **a**-

cells expressing hen egg white lysozyme. This antibody-antigen interaction should promote the formation of cell aggregates (agglutination) and eventually diploid production. The affinity constant for the scFv D1.3/lysozyme interaction [K_a =3 x 10^8 M^{-1} (Hawkins *et al.*, 1993)] is similar to the **a**-/α-agglutinin interaction [K_a = 1 x 10^9 M^{-1} (Lipke *et al.*, 1987)] therefore should mimic the agglutinin's function. The mutants, CAY1 (*sag1-Δ976::kanMX4*) and CAY3 (*aga2::kanMX4*) were also mated under the same conditions to ensure that the agglutinins had been properly disrupted.

Overexpression of vectors, pYES2-*SAG1*-HEWL and pYES2-*SAG1*-scFv D1.3, on the surfaces of the **a**- and α-disrupted strains, respectively, was also investigated.

6.2 MATERIALS

6.2.1 Reagents

Hen egg white lysozyme, α-factor and *Micrococcus lysodeikitus* cells were purchased from Sigma (Product Numbers - L2879, T6901 and M3770, respectively). An Immunoprobe™ biotinylation kit, dialysis tubing (Product No - D0405), phosphocellulose (Product No - C2383) and FAST™ p-nitrophenyl phosphate (pNPP) substrate tablet set were also obtained from Sigma. Streptavidin-alkaline phosphatase (Product No - RPN1234) was purchased from Amersham International and the 0.22μM GS filter papers were obtained from Millipore.

6.2.2 Strains

Saccharomyces cerevisiae strain D665-1A (*MAT***a**) was obtained from the Yeast Genetics Course (1996), Cold Spring Harbor Laboratory, New York.

Note: The parental strains used in this germinal centre strategy (FY10 and FY22) originated from strain S288C (Winston *et al.*, 1995) of which X2180-1A and X2180-1B are derivatives (Mortimer and Johnston, 1986). Strains X2180-1A and X2180-1B have been used in many principal agglutination

studies, hence their agglutinin expression levels and characteristics have been extremely well studied (Terrance *et al.*, 1987; Lipke *et al.*, 1987; Hauser and Tanner, 1989; Lipke *et al.*, 1989).

6.3 METHODS

6.3.1 Estimation of protein concentration

Protein concentrations were estimated by absorbance measurements at 280nm. The molar extinction coefficient used for hen egg white lysozyme was 37932 $M^{-1}cm^{-1}$ (Sophianopoulos *et al.*, 1962).

6.3.2 Functional detection of scFvD1.3 in *E. coli* (EMBO practical course, Pomezia, 1991)

Two individual TOP10F' *E. coli* colonies carrying the vector pUC119His6mycXba-scFv D1.3, were picked from an LB plate (supplemented with 100μg/ml ampicillin) and resuspended in 3ml of XL medium (16g/l Bacto-tryptone, 16g/l Bacto-yeast extract, 5g/l sodium chloride) containing 1% (w/v) glucose and 100μg/ml ampicillin. These cultures were grown overnight at 37°C and 100μl of each overnight culture was transferred to 2ml of XL medium supplemented with 0.1% (w/v) glucose and 100μg/ml ampicillin. The new cultures were grown at 37°C until an absorbance (at 600nm) of 0.9 was reached. Isopropyl-β-D-thiogalactoside (IPTG) was added to a final concentration of 1mM and the cultures were shaken (250 r.p.m.) at 30°C for a further 24 hours. After this time, the cultures were centrifuged at 4000 r.p.m. for 10 minutes and the supernatant from each sample was analysed for functional scFv D1.3 using an ELISA. The supernatant from an uninduced culture in XL medium (supplemented with 100μg/ml ampicillin/1% (w/v) glucose) was used throughout as a negative control.

A hen egg white lysozyme solution at a concentration of 5.5×10^{-5} M was used to coat the wells of an ELISA plate and the protocol described in Section 2.3.4 was performed. A primary polyclonal antibody serum specific

for mouse IgG Fab fragments (rabbit anti-mouse IgG $F(ab')_2$ fragment specific; Table 2.1) was used. Binding of this primary antibody was detected using a horseradish-peroxidase conjugated goat anti-rabbit IgG secondary antibody (Table 2.2).

6.3.3 Overexpression of scFv D1.3 and HEWL on the surface of the disrupted agglutinin mutants

The disrupted α-strain, CAY1 (Table 5.11), containing the overexpression vector, pYES2-*SAG1*-scFv D1.3 was induced with galactose for 24 hours (Section 4.3.3). The disrupted **a**-strain, CAY3 (Table 5.11), carrying pYES2-*SAG1*-HEWL was also treated with galactose and both sets of cells were employed as positive controls in all experiments concerning the expression of lysozyme and scFv D1.3 on the surface of the engineered cells.

6.3.4 Partial purification of a-factor (Betz *et al.*, 1977)

Phosphocellulose (10g) was resuspended in 400ml of water and allowed to swell. The water was decanted and 1 litre of 0.5M sodium hydroxide/sodium chloride was added for 5 minutes. This solution was decanted and the phosphocellulose was washed with 500ml of water using a Büchner funnel (Whatman No1 filter paper). The phosphocellulose was then resuspended in 1 litre of 0.5M hydrochloric acid for 5 minutes. This solution was decanted and the washing step was repeated. The phosphocellulose was then resuspended in 75ml of 0.05M potassium phosphate buffer, pH 6.0 and allowed to settle. The unsettled particles were removed by decanting and this process was repeated several times. A 1ml column (1 by 10cm) of phosphocellulose was prepared and equilibrated with 50ml of 0.05M potassium phosphate buffer, pH 6.0.

The **a**-strain (D665-1A) was used to inoculate 50ml of minimal glucose medium (6.7g/l Bacto-yeast nitrogen base without amino acids, 20g/l glucose) and was shaken (250 r.p.m.) at 30°C for 24 hours. The cells were removed by centrifugation at 3000 r.p.m. (1500g) for 10 minutes using a Sorvall GLC-4 general laboratory centrifuge and the pH of the supernatant

was checked (~pH 2.6). The cell-free supernatant was applied directly to the phosphocellulose column, washed with 1.5ml of 0.05M potassium phosphate buffer, pH 6.0 and eluted with 10ml of 0.5M potassium phosphate buffer, pH 6.0. Fractions (1ml) were collected, adjusted to pH 5.0 and 10% (w/v) glycerol/1mM EDTA was added. These fractions were then stored at -20°C until required.

The activity was analysed by dropping 10µl of each fraction onto BBMB plates that had been previously spread with 5 x 10^5 α-cells (FY10). [The BBMB plates were prepared by adding 0.03% (w/v) methylene blue and 0.1M phosphate-citrate buffer (pH4.5) to molten YPD/agar medium]. The BBMB plates were then incubated at room temperature for 2 days and analysed for the presence of halos (**a**-factor inhibits the growth of the α-cell lawn and therefore the appearance of halos can be used to measure **a**-factor activity). The lowest concentration of **a**-factor that gives a detectable halo is defined as 1 unit of pheromone/ml (Sprague, 1991). The concentration of pheromone present in different fractions was measured by performing dilutions (1 in 2, 1 in 5, 1 in 10) and applying 10µl of these dilutions to a BBMB plate previously spread with a known amount (5 x 10^5 cells) of α-cells.

6.3.5 Induction and detection of scFv D1.3 on the surface of engineered α-cells

The engineered α-strain containing the scFv D1.3 gene fused to the C-terminal end of the α-agglutinin gene (CAY2; *MATα ura3-52 leu2Δ1 sag1-Δ976*::scFv D1.3) was grown to exponential phase (~1 x 10^7 cells/ml) in 10ml of YPD at 30°C. The culture was divided into two parts. Two units of partially purified **a**-factor (Section 6.3.4) were added to one part of the culture (Betz *et al.*, 1978), the other part was left untreated, and both were grown at 30°C for an additional 2 hours. A wild type α-strain (FY10) culture in exponential phase was also prepared for a negative control. The α-agglutinin null mutant, CAY1, carrying the scFv D1.3 overexpression vector was used as a positive control (MATα *ura3-52 leu2Δ1 sag1-Δ976*::*kanMX4* [pYES2-*SAG1*-scFv D1.3]).

After this time, approximately 2 x 10^7 cells from each culture were analysed for scFv D1.3 expression using an enzyme immunoassay (Section 2.3.5). The primary antibody, rabbit anti-mouse IgG $F(ab')_2$ fragment specific (Table 2.1), followed by the secondary horseradish-peroxidase conjugated goat anti-rabbit IgG antibody (Table 2.2) were used for the detection of scFv D1.3. Cycloheximide at a concentration of 10μg/ml was used in all solutions throughout the enzyme immunoassay to stop new protein synthesis.

6.3.6 Biotinylation of HEWL

A) Dialysis of lysozyme

The dialysis tubing was boiled for 10 minutes in 300ml of 2% (w/v) sodium bicarbonate/1mM EDTA (pH 8.0) and rinsed thoroughly with distilled water. The tubing was then boiled for 10 minutes in 300ml of 1mM EDTA, pH 8.0 and the rinsing step was repeated. The tubing was then washed with 0.1M sodium phosphate, pH 7.2 and 10ml of lysozyme solution (5.5 x 10^{-4} M) was placed inside the dialysis tubing. The lysozyme was dialysed in 250ml of 0.1M sodium phosphate, pH 7.2 overnight on a magnetic stirrer, at 4°C.

B) Biotinylation of lysozyme

The dialysed lysozyme solution was biotinylated using the Immunoprobe™ biotinylation kit. A molar ratio of biotinamidocaproate-N-hydroxyl-sulfosuccinimide ester (BAC-SulfoNHS) to HEWL of 5:1 was prepared and the labelling reaction was performed according to the manufacturer's instructions. Biotinylation of the HEWL was confirmed using an ELISA (Section 2.3.4). The wells of an ELISA plate were coated with the biotinylated HEWL at a concentration of 7nM and streptavidin-alkaline phosphatase, diluted (1:1000) in 0.5% (w/v) bovine serum albumin/PBS was used exclusively for detection. Bovine serum albumin (BSA) instead of non-fat dried milk (Marvel) was used as a blocking agent throughout. Alkaline phosphatase activity was analysed using 200μl of FAST™ pNPP substrate in each reaction. The reactions were incubated in the dark at room temperature (for approximately 0.5-2hours) until a yellow colour

developed. The reaction was then stopped with 50µl of 3M sodium hydroxide. The absorbances (at 405nm) were read using a Multiscan® MCC Labsystems ELISA plate reader. After verification, the biotinylated HEWL was lyophilised using a Lyoprep-300 freeze drier at less than 0.05mbar overnight and stored at -20°C.

6.3.7 Functional identification of scFv D1.3 on the surface of engineered α-cells

The engineered α-strain containing scFv D1.3 DNA fused to the C-terminal end of α-agglutinin gene (CAY2; *MATα ura3-52 leu2Δ1 sag1-Δ976*::scFvD1.3) was grown overnight in 10ml of YPD at 30°C. Approximately 2 x 10^7 cells were washed three times with PBS and 8nM of biotinylated lysozyme diluted in 0.5% (w/v) BSA/PBS was added. The cells were vortexed and incubated for 2-3 hours at room temperature on a vertically rotating platform. The cells were collected via centrifugation at 13000 r.p.m. for 30 seconds. The cell pellets were washed three times with PBS/0.05% (w/v) Tween 20 and then three times with PBS. Streptavidin-alkaline phosphatase (1:1000) diluted in PBS/0.05% (w/v) BSA was added to each sample, vortexed and incubated at room temperature for 1 hour on the rotating platform. The washes were then repeated and 200µl of FAST™ pNPP substrate was added to each. The reactions were incubated at room temperature in the dark until a yellow colour developed and stopped with 50µl of 3M sodium hydroxide. The reaction mixtures were re-centrifuged and the supernatants were transferred to an ELISA plate. This protocol was also carried out with a negative control (FY10 only) and a positive control (MATα *ura3-52 leu2Δ1 sag1-Δ976*::*kanMX4* [pYES2-*SAG1*-scFv D1.3]).

6.3.8 Induction and detection of HEWL on the surface of engineered a-cells

The engineered **a**-strain containing hen egg white lysozyme gene fused to the C-terminus of the **a**-agglutinin binding domain (CAY4; *MAT***a** ura3-52 his3Δ200 *aga2*::HEWL) was grown to exponential phase (~1 x 10^7 cells/ml)

in 10ml of YPD. α-Factor was added to 5ml of this culture to a final concentration of 5μM (Sprague, 1991). The other 5ml of the culture was not treated with pheromone and both were left at 30°C for 1 hour at 250 r.p.m. The parental **a**-strain (FY22) and disrupted **a**-strain, CAY3, containing the lysozyme overexpression vector (*MATa ura3-52 his3Δ200 aga2::kan*MX4 [pYES2-*SAG1*-HEWL]) were also grown to exponential phase and were used for negative and positive controls, respectively. Approximately 2 x 10^7 cells from each culture were analysed for lysozyme expression using an enzyme immunoassay (Section 2.3.5). The rabbit anti-HEWL primary antibody (Table 2.1) followed by the horseradish peroxidase conjugated goat anti-rabbit (Table 2.2) secondary antibody were used for detection of lysozyme on the surface of yeast. Cycloheximide at a concentration of 10μg/ml was used throughout the enzyme immunoassay protocol.

6.3.9 Functional identification of HEWL on the surface of engineered a-cells

The disrupted **a**-strain, CAY3, containing the lysozyme overexpression vector (*MATa ura3-52 his3Δ200 aga2::kanMX4* [pYES2-*SAG1*-HEWL]) was patched onto two Synthetic Complete plates lacking uracil. These plates contained different carbon sources, either 2% (w/v) glucose or 2% (w/v) galactose, and were incubated for 3-4 days at 30°C. After this time, both plates were replica-plated onto one lysoplate and incubated at room temperature for several days. The plates were then observed for zones of clearing due to bacterial lysis. The size of the zones was proportional to the concentration of lysozyme (Osserman and Lawlor, 1966).

Lysoplates were prepared by resuspending *M. lysodeikitus* cells in a small amount (1-2ml) of 55mM potassium phosphate buffer, pH 6.2 (Mörsky, 1983) using a vertically rotating platform for 15 minutes. These cells were added to molten 1% (w/v) low melting point agarose in 55mM potassium phosphate buffer, pH 6.2 to a final concentration of 0.5% (w/v) and the cell/agarose suspension was poured into 9cm petri dishes to a depth of 3-4mm (Osserman and Lawlor, 1966).

The **a**-strain containing the hen egg white lysozyme gene fused to the 3' terminus of the **a**-agglutinin binding domain (CAY4; *MAT**a** ura3-52 his3Δ200 aga2::*HEWL) was also examined for functional lysozyme on its cell surfaces using the lysoplate method. This strain and the wild type **a**-strain (FY22) were patched onto the same YPD plate and incubated at 30°C for 1-2 days. The plate was then replica plated onto a lysoplate, incubated for several days and analysed for zones of clearing around the replicated patches.

6.3.10 Agglutination assay (Lipke *et al.*, 1989)

Three millilitres of YPD in 15ml centrifuge tubes were inoculated with approximately 10^7 cells of each of the mating types to be tested. The tubes were vortexed and incubated vertically overnight at room temperature on a rotating platform (120 r.p.m.). Agglutination should result in a lacy cell pellet that spreads out over the bottom of the tube, whereas non-agglutinated mixtures should produce a compact cell pellet.

6.3.11 Solid mating assay (Lipke *et al.*, 1989)

Approximately 2 x 10^7 mid-log phase cells of each mating type were mixed in a 1.5ml Eppendorf tube. The cells were collected on a filter paper (0.22μM GS) using a Büchner funnel and the filter was placed on a YPD plate for 6 hours at 30°C. The filter was then resuspended in 30ml of 1M sorbitol in a 50ml centrifuge tube and vortexed vigorously. The cells were diluted appropriately and plated onto both non-selective medium (YPD) and Selective Complete medium lacking leucine and histidine, which selected for diploids. These plates were incubated for 2 days at 30°C and the resulting colonies were counted on both the selective and non-selective plates. The mating frequencies were determined by dividing the number of colonies on selective medium (diploids) by the number of colonies on the non-selective medium (diploids and haploids).

6.3.12 Liquid mating assay (Lipke *et al.*, 1989)

Approximately 2 x 10^7 mid-log phase cells of each mating type were mixed in 5ml of YPD medium in a 50ml conical flask and incubated at 30°C with rotation (300 r.p.m.) for 6 hours. The cells were vortexed vigorously, diluted and plated onto both non-selective medium (YPD) and Selective Complete medium lacking leucine and histidine. These plates were incubated for 2 days at 30°C and the mating frequencies were calculated as described in Section 6.3.11.

6.4 RESULTS

6.4.1 Functional expression of scFv D1.3 in *E. coli*

E. coli cells containing vector pUC119His6mycXba-scFv D1.3 were shown to be expressing functional scFv D1.3 antibody fragments after an induction time of 24 hours. The supernatant of the uninduced culture displayed no lysozyme binding activity (Table 6.1).

Table 6.1 - ELISA results obtained from the supernatants of induced and uninduced cultures of TOP10F' [pUC119His6mycXba-scFv D1.3].

Plasmid	IPTG	Abs_{450}[a]
pUC119His6mycXba-scFv D1.3	+	0.50, 0.58, 0.53 (0.54)
"	+	0.54, 0.55, 0.46 (0.52)
"	-	0.05, 0.02, 0.01 (0.03)

[a] Values in the brackets represent the mean of the three individual determinations.

6.4.2 Partial purification of a-factor

a-factor is not commercially available because it is difficult to synthesize due to its hydrophobic nature, structure and associated farnesyl group (Proteau *et al.*, 1990). Therefore it was partially purified from the culture supernatant of a MATa strain (D665-1A) using ion-exchange chromatography. The chosen a-strain contained no auxotrophic mutations due to the fact unmanipulated strains are known to produce a large amount of pheromone (Betz *et al.*, 1977). The fractions eluted from the phosphocellulose column were analysed for a-factor activity using BBMB plates. All a-factor activity was present in the first two millilitres of the eluted fractions and the amount of activity was determined semiquantitatively by analysing serial dilutions. A total of three units of pheromone was recovered which was comparable with results obtained for different a-strains (Betz *et al.*, 1977).

6.4.3 Detection of scFv D1.3 on the surface of engineered α-cells

The engineered α-cells containing the scFv D1.3 gene integrated into the 'N/C' junction of the α-agglutinin gene (strain CAY2) were analysed for cell surface expression of scFv D1.3 antibody fragment in the presence and absence of **a**-factor. Enzyme immunoassay results indicated that the Sag1p-scFv D1.3 fusion was detectable on the surface of **a**-factor treated cells but the surface expression of the untreated cells was not significantly above background (Table 6.2). The α-agglutinin null mutant, CAY1, carrying the scFv D1.3 expression vector was shown to be displaying a large amount of scFv D1.3 on its cell surface. Cycloheximide was used throughout the enzyme immunoassay to stop protein production. Hence, the agglutinin molecules that were analysed had become accessible at the cell surface directly after the pheromone treatment or galactose induction (Betz *et al.*, 1978).

Table 6.2 - Enzyme immunoassay results for the detection of scFv D1.3 antibody fragments on the surface of pheromone-treated and untreated cells.

Strain[a]	a-factor	Abs_{450}[b]
CAY2	+	0.45, 0.47, 0.48 (0.47)
CAY2	-	0.33, 0.45, 0.49 (0.42)
CAY1 [pYES2-*SAG1*-scFv D1.3]	-	1.41, 1.41, 1.48 (1.43)
Wild type (FY10)	-	0.32, 0.35, 0.39 (0.35)

[a] The genotypes of the strains used in this experiment are summarized in Table 5.11.

[b] Values in the brackets represent the mean of the three individual determinations.

6.4.4 Functional identification of scFv D1.3 using biotinylated lysozyme

Lyzoyme was dialysed overnight and biotinylated using the Immunoprobe™ biotinylation kit. Biotinylation was confirmed by coating an ELISA plate with the biotinylated lysozyme and screening with streptavidin-

alkaline phophatase. A mean absorbance (at 405nm) of 0.48 for three biotinylated samples compared to 0.05 for three non-biotinylated samples was obtained.

The engineered α-cells, CAY2, and α-agglutinin disrupted strain, CAY1, overexpressing scFv D1.3 were analysed for their ability to bind to lysozyme using the avidin-biotin system. Bound lysozyme was detected using streptavidin-alkaline phosphatase and absorbances obtained for both sets of cells were reproducibly above background (Table 6.3). Unfortunately, the ten-fold increase expected for the cells overexpressing the scFv D1.3 was not observed. These results indicated that there was some level of functional scFv D1.3 on the cell surfaces of the engineered α-strains.

Table 6.3 - Absorbance values obtained using the avidin-biotin system for the identification of functional scFv D1.3 on yeast cell surfaces.

Strain[a]	Abs_{405}[b]
CAY2	0.13, 0.14, 0.14 (0.14)
CAY1 [pYES2-*SAG1*-scFv D1.3]	0.17, 0.17, 0.17 (0.17)
Wild type (FY10)	0.05, 0.07, 0.07 (0.06)

[a] The genotypes of the strains used in this experiment are summarized in Table 5.11.

[b] Values in the brackets represent the mean of the three individual determinations.

6.4.5 Detection of HEWL on the surface of engineered a-cells

The engineered **a**-cells containing the hen egg white lysozyme gene fused to the 3' end of the **a**-agglutinin binding domain (strain CAY4) were examined for the expression of lysozyme on their surfaces in the presence and absence of α-factor. Enzyme immunoassay results indicated that lysozyme was detectable on the surface of both pheromone-treated and untreated **a**-cells. The induced null **a**-agglutinin mutant, CAY3, containing the lysozyme expression vector was shown to be displaying a large

amount of lysozyme on its cell surface compared to the engineered **a**-strain (Table 6.4).

Table 6.4 - Enzyme immunoassay results for detection of lysozyme on the surface of pheromone-treated and untreated cells.

Strain[a]	α-factor	Abs_{450}[b]
CAY4	+	0.28, 0.32, 0.34 (0.31)
CAY4	-	0.25, 0.26, 0.26 (0.26)
CAY3 [pYES2-*SAG1*-HEWL]	-	1.75, 1.84, 2.61 (2.07)
Wild type (FY10)	-	0.14, 0.14, 0.15 (0.14)

[a] The genotypes of the strains used in this experiment are summarized in Table 5.11.

[b] Values in the brackets represent the mean of the three individual determinations.

6.4.6 Verification of functional HEWL using *M. lysodeikitus* cells

The expressed lysozyme on the surface of the engineered **a**-cells (CAY4) was analysed for activity using *M. lysodeikitus* cells. No zones of clearing due to bacterial lysis were observed around the engineered **a**-cells after several days of incubation (Figure 6.1a). Therefore, it was presumed the expression level of lysozyme on the surface of the engineered **a**-cells was lower than the minimal amount detectable by the lysoplate method [of the order of μg/ml (Osserman and Lawlor, 1966)]. A zone of clearing was observed around the cells overexpresssing the lysozyme (Figure 6.1b) thereby demonstrating the lysozyme was active.

Figure 6.1 - Lysoplate test for detection of lysozyme activity.

a) Wild type **a**-strain, FY22. (*MATa ura3-52 his3Δ200*)

Engineered **a**-strain, CAY4. (*MATa ura3-52 his3Δ200 aga2*::HEWL)

b) Induced and uninduced patches of the **a**-agglutinin null mutant, CAY3, containing the lysozyme expression vector (*MATa ura3-52 his3Δ200 aga2*::*kanMX4* [pYES2-*SAG1*-HEWL]).

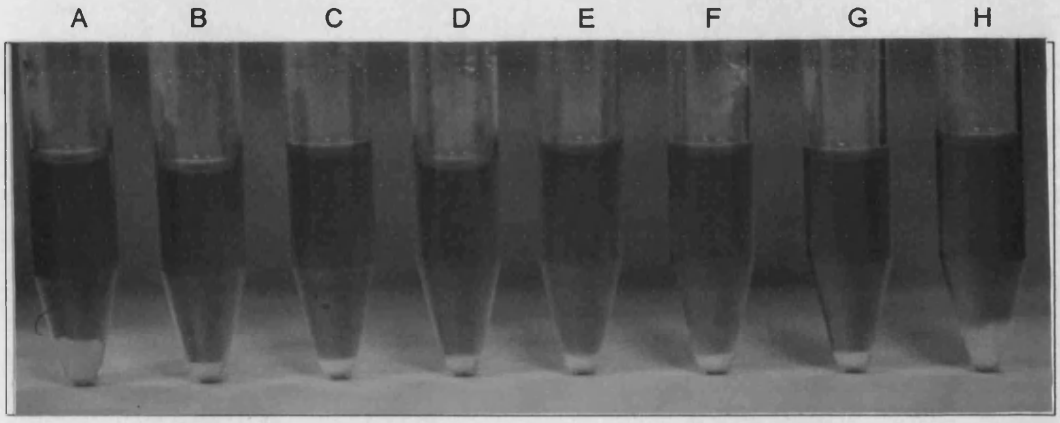
6.4.7 Agglutination results

Each cross described in Figure 6.2 was tested for agglutination. The parental strain cross (FY10 x FY22) produced the expected lacy cell pellet at the bottom of the tube (Lanes A and H). When both the **a**- and α-agglutinin disrupted mutants (CAY3 and CAY1, respectively) were mixed with the opposite parental mating type strains, no cell aggregates were detected. This confirmed the *kanMX4* cassette had disturbed the agglutinin function at both loci.

Unfortunately, the crossing of the α-strain expressing the integrated scFv D1.3 (CAY2) with the **a**-strain expressing the integrated lysozyme

(CAY4) did not produce the desired agglutination result (Lanes D and E). A compact cell pellet (no cell aggregates) was observed which indicated that the lysozyme-scFv D1.3 molecules on the surface of the relevant mating types were not interacting sufficiently to promote cell to cell contact.

Figure 6.2 - Agglutination test



Lanes	Mating type crosses[a]
A and H	FY10 x FY22
B and C	CAY1 x FY22
D and E	CAY2 x CAY4
F and G	FY10 x CAY3

[a] = The genotypes of these strains are described in Table 5.11.

6.4.8 Mating results

When the mating experiments were performed on solid media, there was a slight reduction (four-fold) in the mating frequencies of the agglutinin-disrupted mutants compared to the parental strains. The same crosses were performed in liquid media and characteristically the mating frequencies for the null mutants were four orders of magnitude lower than for the wild type strains (Table 6.5).

The engineered **a**- and α-cells expressing Aga2p-HEWL and Sag1p-scFv D1.3 fusions on their cell surfaces, respectively, also were mated in both liquid and solid media. Unfortunately, the mating frequencies obtained for both the liquid and solid media were similar to the disrupted mutants. This result indicated that the surface lysozyme and scFv D1.3 molecules were not interacting sufficiently in courtship to promote the stable cell-to-cell contact required for diploid formation.

Table 6.5 - Mating frequencies [a] obtained for the engineered and mutant **a**- and α-cells.

Cross[b]	Solid medium	Liquid medium
FY10 x FY22	1, 1	1, 1
CAY1 X FY22	0.21, 0.27	4.0×10^{-4}, 4.9×10^{-4}
FY10 X CAY3	0.20, 0.25	2.5×10^{-4}, 4.0×10^{-4}
CAY2 X CAY4	0.21, 0.25	3.2×10^{-4}, 7.3×10^{-5}

[a] Relative to the wild-type frequency. Each cross was performed in duplicate.

[b] The genotypes of the strains used in this experiment are summarized in Table 5.11.

6.5 DISCUSSION

6.5.1 Expression levels

The scFv D1.3 antibody fragment was exported into the bacterial supernatant and was shown to be binding specifically to the antigen, lysozyme (Table 6.1).

The surface expression levels for both the Aga2p-HEWL and Sag1p-scFv D1.3 fusions on the relevant engineered cells appeared to be very low. Firstly, the expression level for the Sag1p-scFvD1.3 fusion on the surface of the engineered α-strain was only detectable by an enzyme immunoassay after treatment with partially purified **a**-factor (Table 6.2). It was presumed that some level of Sag1p-scFv D1.3 expression would be identified on the

surface of the untreated engineered α-cells due to the fact that approximately 5 x 10^4 agglutinin molecules are constitutively expressed (Wojciechowicz and Lipke, 1989). The α-strain (X2180-1B) used to calculate the number of surface agglutinin molecules is a relative of the α-strain (FY10) used in this project and therefore should have similar agglutinin expression levels. Wojciechowicz and Lipke (1989) also observed a 1.3 fold increase in the α-agglutinin expression level of **a**-factor treated α-cells (X2180-1B) using an enzyme immunoassay, which is comparable to the 1.7 fold increase detected in these experiments.

Results obtained using the biotinylated lysozyme indicated that a low level of functional scFv D1.3 molecules was present on the surface of the engineered α-cells (Table 6.3). Surprisingly, the biotinylation results acquired for the cells overexpressing scFv D1.3 on their surfaces were not the expected order of magnitude higher. The limited amount of biotinylated lysozyme used in these experiments may have affected these results. Consequently, performing experiments with different amounts of biotinylated lysozyme may have been more definitive.

A 1.4 fold increase in the expression of the Aga2p-HEWL fusion on the surface of engineered **a**-cells was observed using an enzyme immunoassay after treatment with α-factor (Table 6.4). This is comparable to the 2-2.5 fold increase obtained for the native **a**-aggluitnin using the α-factor treated **a**-strain, X2180-1A (Terrance and Lipke, 1987). The level of lysozyme expression was too low to show functionality using the lysoplate method (Figure 6.1a) and a more sensitive method such as a turbidimetric assay (Kato *et al.*, 1994b) could possibly be applied in future to calculate the amount of enzymatic activity on the surface of these cells. A preliminary turbidimetric assay was performed with the engineered **a**-cells but the results obtained were found to be unreproducible. Lysozyme overexpressed on the disrupted **a**-cells did cause bacterial lysis indicating that it was active on the cell surface (Figure 6.1b). Functional detection of the lysozyme in this experiment could have been assisted by the higher level of expression causing diffusion of some of the lysozyme molecules into the surrounding agar. Experiments analysing the expression of the scFv B7 antibody

fragment on the surface of yeast cells indicated that a certain amount of the antibody fragment was secreted into the culture supernatant (Section 4.4.4b).

With the limited information available, the expression levels for both the Aga2p-HEWL and Sag1p-scFvD1.3 fusions appear to be low. The reasons for these low expression levels could be numerous. Due to the fact that scFv D1.3 and lyzozyme are foreign to *S. cerevisiae,* elements such as codon usage adversely affecting the stability of the mRNA and fortuitous sequences causing pauses in transcription may lead to lower expression levels (Romanos *et al.* 1992). Furthermore, the *loxP 511* sites in the middle of the protein fusions may be affecting expression either by being more susceptible to proteolytic cleavage (Alfthan *et al.*, 1995) or by the incorrect initiation of transcription due to the two TATA sequences. Expression of the Sag1p-scFvD1.3 fusion could also have been hindered by the position of *loxP 511* sequence between the promoter and initiation site affecting transcription. Additionally, palindromic sequences such as *loxP* sites have been reported to form secondary structures, which can drastically inhibit translation (Romanos *et al.*, 1992).

Sequencing results from Chapter 5 showed that both the Aga2p-HEWL and Sag1p-scFv D1.3 fusions had the correct molecular information (i.e. signal sequences, GPI anchor attachment sequences, etc.) to be expressed through the secretory pathway and become incorporated on the cell surface (Figures 5.9 and 5.14). The Sag1p-scFvD1.3 protein fusion was shown to be functional on the surface of the engineered α-cells, therefore confirming the fusion was in the correct reading frame. Further experiments are required to show the Aga2p-HEWL fusion is active on the surface of the engineered **a**-cells. Additionally, experiments are required to prove that the agglutinin fusion proteins are covalently attached to the β-1,6-glucan component of the yeast cell wall.

6.5.2 Attempted mating

The unusual phenotype of a small decrease in the mating frequency on solid media and a three to five-fold reduction in liquid media is a characteristic of

agglutinin mutants (Lipke *et al.*, 1989; Roy *et al.*, 1991). Both the constructed α- and **a**-agglutinin mutants, CAY1 and CAY3, displayed this phenotype (Table 6.5). Expression of aminoglycoside phosphotransferase, which is encoded by the *kanMX4* cassette, does not affect mating frequency (Hadfield, 1990). Therefore, it can be presumed that the fusion of the *kanMX4* cassette to the C-terminus of the **a**-agglutinin binding domain (*AGA2*) and the replacement of the N-terminal binding domain of the α-agglutinin (*SAG1*) with the *kanMX4* cassette disrupted the agglutinin functions.

Theoretically, the replacement of the selectable marker, *kanMX4*, with the antigen and antibody DNA should mimic the **a**-/α-agglutinin interaction and promote the stable cell-to-cell contact necessary for mating. Unfortunately, the mating of the engineered α- and **a**-strains expressing the Sag1p-scFv D1.3 and Aga2p-HEWL fusions, respectively, did not produce the desired results. No cell aggregates were formed in the agglutination experiments (Figure 6.2) and no increase in mating frequencies compared to the agglutinin mutants was observed (Table 6.5). The agglutinability of the engineered and mutant strains was not quantified. Subtle differences between these cell types may have been detected by calculating an agglutination index for each mating cross (Terrance and Lipke, 1981). The agglutination index (AI) can be calculated by measuring the formation of cell aggregates in a mating reaction using a spectrophotometer [AI= 1 - (2 x $A_{\alpha^{+a}}$)/(A^a +A_α); where A is the mean optical density (at 600nm) obtained for the cell types or type indicated by the superscript]. The agglutination index varies from 0 to 1, with higher indices denoting a greater degree of agglutination.

The factors affecting the desired scFv D1.3-lysozyme interaction between the engineered α- and **a**-cells could be numerous. It can be presumed that the scFv D1.3 and lysozyme proteins were accessible on the exterior surface of the mating cells due to the fact the reagents used were able to detect them. The strength of the agglutination reaction in *S. cerevisiae* is affected more by the number of binding sites than by the binding constant (Terrance and Lipke, 1987). The expression results

indicated that only low levels of scFv D1.3 and lysozyme were present on the surface of the engineered cells and consequently the number of binding sites available may have been the major limiting factor in promoting cell-to-cell contact. Hyperglycosylation of scFv D1.3 and/or HEWL proteins may also have inhibited their interaction (Romanos *et al.* 1992).

a- and α-cells have been shown to discriminate between different levels of pheromone and during courtship partners which are producing the highest level of pheromone are selected (Jackson and Hartwell, 1990). Lipke and Kurjan (1992) suggested the agglutinins might have a subsidiary role in increasing the local pheromone concentration at the shmoo tips of mating cells, especially in liquid media where pheromone concentration is affected by diffusion. Therefore, perhaps the scFv D1.3 and HEWL proteins were unable to increase the local pheromone concentration by this unknown mechanism which may also explain the observed mating defects of the engineered strains.

6.5.3 Future work

It is unknown whether the maximum level of expression was achieved for both the pheromone-treated **a**- and α-engineered cells, although the suggested amount of pheromone required to achieve maximum expression was used in both cases (Betz *et al.*, 1978; Terrance and Lipke, 1981; Guthrie and Fink, 1991). Further experiments investigating pheromone induction time and concentration would help resolve this issue. A time course monitoring physiological changes such as shmoo formation or arrest in G1 phase might give the most favourable conditions. Replacing the scFv D1.3 or HEWL DNA with a reporter gene (i.e. β-galactosidase or green fluorescent protein) could be another way of estimating the maximum level of expression directly. In fact, green fluorescent protein (GFP) was integrated into the α-agglutinin locus using the Cre-*lox* recombination system but initial expression experiments were unable to locate the GFP on the surface of the engineered α-cells (personal communication, Mark Redfern, this laboratory). Additionally, using **a**-cells which lack the barrier protease which degrades α-factor (*bar1* mutants) may have increased the level of

Aga2p-HEWL expression (Mackay *et al.*, 1988). Although the expression of Sag1p-scFv D1.3 and Aga2p-HEWL on the surface of the pheromone-treated cells was not optimised in the individual detection experiments, the pheromone concentration presumably would have been optimal in the mating experiments.

Ideally, the number of molecules of scFv D1.3 and lysozyme on the surface of the engineered cells should have been calculated and compared to the display efficiency of the native agglutinins. Binding assays using variable amounts of labelled lysozyme or scFv D1.3 with a known amount of the appropriate engineered cells expressing scFv D1.3 or lysozyme on their cell surfaces would give the information required to calculate this (Terrance *et al.*, 1987). Additionally, two-colour flow cytometry, which has been used to estimate the display efficiency of the Aga2p-scFv fusion (under the control of a *GAL1* promoter) on the surface of **a**-cells (Boder and Wittrup, 1997), may also have proved informative.

6.5.4 Conclusions

Both the α- and **a**-null mutants, CAY1 and CAY3, were shown to have their agglutinin functions disrupted by the selectable marker (*kanMX4*). Unfortunately, the strategy to mimic the germinal centre reaction by mating the engineered α- and **a**-strains, CAY2 and CAY4, was not achieved. The engineered **a**- and α-cells were shown to be expressing the antigen, lysozyme and the antibody, scFv D1.3 (analogous to follicular dendritic cells and B-cells), respectively. Crossing of these engineered strains did not increase the mating frequency compared to the agglutinin null mutants. These results indicate the scFv D1.3 and lysozyme were not interacting sufficiently to promote the stable cell-to-cell contact needed for the formation of diploids.

CHAPTER 7

GENERAL DISCUSSION

The overall aim of this project was to develop strategies for producing high affinity monoclonal human antibodies using the budding yeast, *Saccharomyces cerevisiae*.

In vivo, the germline antibody DNA is subjected to random mutation and antibodies with improved affinity for antigen are selected (a process known as affinity maturation). This stepwise process occurs in the germinal centre (Section 1.3.3) and the ultimate objective of this project was to generate high affinity human antibodies by copying this system. The current methods available such as hybridoma technology (Section 1.4.1), humanization (Section 1.4.2) and phage display (Section 1.4.4) are moderately slow at producing human monoclonal antibodies and are often unable to produce the quality required for therapy. At present, no antibody improvement scheme exists where both the mutation and antigen-selection processes can be controlled within a cell. Eventually, it was intended to create a fast and effective 'in-cell' mutational and antigen-selection yeast scheme. This scheme was devised to overcome all the extensive DNA manipulations that are required for present antibody improvement techniques. Although, this ultimate aim was not accomplished, the work described in this project has laid the foundations to achieve this goal.

7.1 Mimicking the genetic display package of B-cells

B-cells display antibodies on their cell surfaces and in this project an *S. cerevisiae* expression system was constructed which imitated this characteristic (Part I; Chapters 3 and 4). This system utilized the molecular information from the signal sequence and C-terminal anchorage domain of the α-agglutinin gene (Section 1.7) to target and covalently attach antibody fragments to the yeast cell wall. Single chain Fv antibody fragments were expressed on the surface of *S. cerevisiae* (Section 4.4.4c and 4.4.4e) and

shown to be fully functional (Section 4.4.4f). The utilization of the α-agglutinin cell wall protein to display antibodies on the cell surface is similar in certain respects to phage display technology (Section 1.4.4).

The minor and major phage coat proteins have been used extensively to display antibodies on the surface of phage. Although phage display libraries have been effectively employed to isolate a large number of different antibodies (reviewed in Winter *et al.*, 1994; Winter, 1998), unpredictable expression biases exist. Random sampling of a phage display library expressing scFv antibody fragments indicated that half the library sequences were undetectable in the culture supernatant (Vaughan *et al.*, 1996). These expression biases are considered to be due to the fact that incorporation of the protein fusions into the phage particle depends upon the ability of *E. coli* to express them in a soluble form. *E. coli* is a prokaryotic microorganism, therefore lacks the endoplasmic reticulum-specific post-translational processing (i.e. disulphide bond formation, etc.) required for the efficient folding of eukaryotic proteins such as antibodies. Even closely related antibody sequences can demonstrate widely different expression characteristics in *E. coli* (Knappik and Plückthun, 1995; Kipriyanov *et al.*, 1997; Nieba *et al.*, 1997). *S. cerevisiae* is an unicellular eukaryote (Section 1.5) and hence should alleviate the expression biases observed in prokaryotic expression systems.

Once surface expression of antibody fragments was established, the next part of this project was to develop a selection procedure that could isolate yeast cells expressing high affinity antibodies. The selection procedure would have to distinguish between cells that are expressing high affinity antibodies from cells displaying moderate affinity antibodies at higher expression levels. Part 1 of this project was designed with the intention of isolating high affinity antibodies directly using fluorescence-activating cell sorting (FACS). Before this work could be carried out, a two colour FACS method was reported to have successfully isolated yeast cells displaying high affinity antibodies from a small surface display antibody library (Boder and Wittrup, 1997; Kieke *et al.*,1997). Yeast cells which were highly fluorescent because of more efficient expression rather than higher affinity for the

fluorescently labelled target were normalized using a second fluorophore which recognised a non-mutated part of the displayed antibody fragment such as an epitope tag. Since the FACS selection method had already been established, Part II of this project was designed for the purpose of selecting high affinity antibodies using a novel yeast mating strategy (discussed in Section 7.3).

7.2 Towards the construction of a mutant antibody library

For the yeast mating strategy to be successful it was deemed necessary to have constant expression levels on the surface of *S. cerevisiae*. Constant surface expression should overcome the selection problem of identifying yeast cells expressing high affinity antibodies from a population of cells displaying moderate affinity antibodies at higher expression levels. It was decided constant surface expression levels could be achieved by integrating the antibody (or antigen) DNA into the appropriate agglutinin loci (Schreuder *et al.*,1996).

To isolate high affinity antibody fragments on the surface of *S. cerevisiae*, a mutant antibody library was eventually to be created. With this intention, the Cre-*lox* site-specific DNA recombination system from bacteriophage P1 (Section 5.1.1) was recruited to integrate the antibody (and antigen) DNA into the yeast genome. The single chain Fv was the chosen antibody format for this stage of the project, mainly due to the success of expressing the scFv on the surface of *S. cerevisiae* in Part I of this project. Additionally, all the genetic information for the scFv antibody fragment is contained within one 'gene' therefore it is easier to manipulate. For some therapeutic applications such as *in vivo* diagnostic imaging, the scFv is the antibody fragment of choice due to its ability to penetrate tumours owing to its small size (Huston *et al.*, 1993).

The Cre-*lox* system was successful at integrating the scFv antibody fragment DNA into the exact position at the α-agglutinin locus but unfortunately it was found that the frequency of undesirable recombination between the *loxP* and *loxP 511* sites at the locus (or on the exchange vector) was quite high. In some clones, the selectable marker was removed from the

genome without the replacement of the antibody DNA (Sections 5.6.5 and 5.6.7). This undesirable recombination will limit the theoretical size of the antibody library. Hypothetically, the size of the library created using the Cre-*lox* system should only be limited by the efficiency of the transformation and a library of 10^6 clones (Schiestl *et al.*, 1993) could be possible. With the results reported in this project, a potential library of approximately 10^4 clones is possible. A library of 10^4 clones is very small but is two orders of magnitude higher than attainable by standard integration methods. Future work investigating the cause of the illegitimate recombination problem would be extremely beneficial and hopefully lead to an increase in this potential library size.

It can be envisaged that a mutated scFv antibody library could be integrated into the α-agglutinin locus using the Cre-*lox* system. Although the 3D structure of the antigen-antibody (lysozyme-scFv D1.3; Figure 5.2) interaction was available for this project, often the exact structures of therapeutic 'lead' antibodies are unknown. Therefore, a random mutagenesis approach, analogous to the human immune system, is particularly favourable for antibody improvement.

The complementarity determining regions (CDRs) which form the antigen-combining site are normally the first area to be targeted by other mutatgenic approaches (Tsumoto *et al.,* 1995; Yang *et al.*, 1995; Li *et al.*, 1996). Often it is difficult to predict which residues in the CDRs should be targeted by site-specific mutagenesis and sometimes mutated amino acids not located at the antigen-antibody interface lead to an improved affinity (Foote and Winter, 1992; Hawkins *et al.*, 1993). These 'framework' mutations can affect affinity indirectly by influencing the position of the side chain contacting the antigen, providing new contact residues or by replacing 'repulsive' or low affinity contact residues with contact residues with more favourable energetics (Hoogenboom, 1997).

The random mutagenesis method recruited in this strategy must be carefully chosen for an appropriate mutation rate and its ability to generate single nucleotide substitutions. A mutation frequency that incorporates one mutation per antibody gene, similar to the mutation rate that is observed

within the germinal centre (Section 1.3.4; Berek and Milstein, 1988), is well suited for the evolution of antibody fragments in a stepwise manner. Mutation frequencies should not be too high otherwise the combination of beneficial and deleterious mutations may result in the formation of an inefficient affinity maturation strategy (Kepler and Perelson, 1993; Oprea and Perelson, 1997). Additionally, mutation rates that are too low would cause a mutant library to be predominately wild type.

Successful phage display affinity maturation strategies using either error-prone PCR (Hawkins *et al.*, 1993) or a bacterial mutator strain (Low *et al.*, 1996) have been reported and these techniques could be used to create a mutant antibody library before integration into the yeast genome. Furthermore, a DNA shuffling method known as 'sexual' PCR (Stemmer, 1994) which involves the random mutagenesis of several positively selected DNA molecules followed by DNA fragmentation and reassembly, could be a very fast way of evolving high affinity antibodies. However, all these methods involve extensive DNA manipulations outside of the cell before and in-between the antigen-selection process. Hopefully, the work described in this project will assist in the development of a total ' *in vivo* ' mutational system, whereby mutation and antigen-selection can occur in a stepwise fashion, without any requirement for external DNA manipulations. Yeast mutator strains do exist which are deficient in DNA repair mechanisms (Vonborstel *et al.*, 1993; Yang *et al.*, 1996). Using a temperature sensitive strain, where the mutational activity can be controlled, could provide an extremely fast and effective affinity maturation strategy.

Once a clone expressing a high affinity antibody is isolated, it can be envisaged that the Cre-*lox* recombination system could also be used to obtain the advantageously mutated antibody DNA for further analysis. An exchange vector comprising of a marker gene flanked by *loxP* and *loxP 511* sites could be transformed into the yeast cells containing the desired mutated antibody DNA. Cre recombinase expression should catalyse the recombination event between the mutated antibody DNA in the genome and the exchange vector. Yeast cells acquiring the auxotrophic resistance encoded by the marker gene would represent the desired exchange event. Total genomic DNA from these yeast cells could then be transformed into *E.*

coli and transformants conveying the same antibiotic resistance as the exchange vector would contain the cloned mutated antibody DNA. Using the Cre-*lox* system in this way could bypass the need for extensive DNA manipulations that are required for recloning of the antibody DNA from the yeast genome.

The novel application of the Cre-*lox* system developed in this project demonstrates considerable potential for integrating any foreign DNA into a particular locus in the *S. cerevisiae* genome. This strategy does not use auxotrophic markers and therefore can be used in any yeast strain. The efficiency of integration using the Cre-*lox* system is two magnitudes higher than standard integration methods and therefore could be used in a number of applications such as functional genomic analysis studies or recombination-activated gene expression.

7.3 Attempting to copy the antigen-selection process of the germinal centre

During the germinal centre reaction (Section 1.3.3), B-lymphocytes (centrocytes) expressing advantageously mutated antibodies bind to antigen-presenting follicular dendritic cells for survival. In the second part of this project (Part II), the antigen-selection process of the germinal centre was copied by attempting to mate engineered **a**- and α-agglutinin haploid mutants. Agglutinins (Section 1.6.2) are essential for mating under conditions that do not promote stable cell to cell contact (liquid media) and this characteristic was exploited in the antigen-selection mating strategy.

Firstly, the centrocytes and follicular dendritic cells of the germinal centre reaction were copied using the α- and **a**-haploid cells containing the appropriate antibody and antigen DNA in the relevant agglutinin loci, respectively. Unfortunately, expression of the antibody and antigen agglutinin fusions on the surface of the engineered cells was very low (Sections 6.4.3 and 6.4.5). Upon mating of these individually engineered haploid cells, in conditions that did not promote cell-to cell-contact, the antibody and antigen proteins did not interact sufficiently to cause agglutination. Therefore, the mating frequency of the engineered haploid cells expressing the antigen and

antibody DNA was comparable to the mating frequency of agglutinin disruption mutants (Section 6.4.8). The reasons why this antigen-selection mating strategy did not work could be numerous.

Presumably, the major limitation of this strategy was the low surface expression obtained for both the antigen and antibody agglutinin fusions. The agglutination reaction depends greatly on the number of agglutinin molecules on the yeast cell surface (Terrance and Lipke, 1987). Therefore, it would be beneficial to determine the number of agglutinin fusions on the surface of the engineered **a**- and α-haploid cells. If, as expected, the number of agglutinins is low, increasing the surface expression of the agglutinin fusions maybe worthwhile. The surface expression of the agglutinin fusions could be increased by integrating multiple copies of the relevant DNA into the genome (Schreuder *et al.*, 1996) or placing a strong inducible promoter upstream from the fusions (Boder and Wittrup, 1997).

Another point of investigation could be to remove the *lox* sites from the promoter (for the scFv/α-agglutinin fusion) and terminator (for the antigen/**a**-agglutinin fusion) regions of the agglutinin loci. The genomic position of these *lox* sites could be affecting the transcription of the agglutinin fusions and hence the overall surface expression levels. The antibody or antigen genes could have been integrated into the relevant agglutinin loci using a two-step gene replacement method (Scherer and Davis, 1979) instead of the Cre-*lox* recombination system. The two-step gene replacement method involves using a yeast integrating plasmid that contains a mutant copy of the allele of interest. Upon integration at the relevant locus duplication of the region will occur. The region of interest will then comprise of one wild type allele and one mutant allele separated by the plasmid sequence. Homologous crossovers between the mutant and wild type alleles result in excision of the plasmid and loss of one of the copies at the duplicated region. Depending on the exact location of the crossover, the copy left behind will be either the mutant or the wild type form of the gene. Southern hybridization, PCR screening and/or phenotype analysis are then required to distinguish between colonies containing either the wild type or mutant form of the gene. Although the efficiency of integration would have been lower using this two-

step gene replacement method the absence of *lox* sites may have increased the surface expression of the antibody and antigen agglutinin fusions. Hopefully, increasing surface expression of the agglutinin fusions will promote stable cell-to-cell contact between the mating cells and produce a successful antigen-selection strategy.

Agglutinin molecules accumulate in the buds of shmooing cells (Watzele *et al.*, 1988), therefore engineered yeast cells should be analysed by immunofluorescence to check if the agglutinin fusions are constantly expressed over the surface of the yeast cell. Accumulation of fusions in the buds of mating cells could affect the ability of the selection system to discriminate between cells expressing antibodies with slightly different affinities since the binding of antigen will be dictated by both affinity and avidity.

Once the problems of mating antigen-selection process are solved, the artificial yeast 'germinal centre' could be established. Using this strategy, the yeast cells that have formed diploids due to a positive antibody-antigen interaction could be sporulated under nutritional starvation conditions. The haploid α-cells containing the positively selected mutated antibody DNA could then undergo further rounds of mutation and selection without any requirement for extensive DNA manipulations. Finally, a yeast cell displaying a very high affinity antibody could be isolated which contains various advantageous point mutations within its antibody DNA.

7.4 General perspective of this project

A strategy that can harness all the aspects of the immune system should have the capability to create antibodies with a wide variety of affinities and specificities. The work described in this project provides the potential to eventually create a novel 'immune' system for the production of high affinity antibodies using the yeast, *S. cerevisiae.*

New therapeutic 'lead' antibodies could be isolated from a *S. cerevisiae* surface display library composed of human unrearranged variable light and heavy gene segments. Using unrearranged genes should avoid the biases created by *in vivo* tolerance mechanisms (Hoogenboom, 1997) and

antibody fragments towards self-antigens such as tumour or disease associated epitopes could be isolated with great therapeutic value. This library would be analogous to a naïve primary antibody repertoire of the human immune system without the self-bias.

Once a 'lead' antibody is isolated, its affinity to a particular therapeutic target could then be improved through a stepwise mutation and antigen-selection 'germinal centre' strategy as described in this project. Eventually, it is hoped that this strategy will produce therapeutic antibodies faster and more effectively than any system previously developed.

REFERENCES

Abbas, A. K., Lichtman, A. H. and Pober, J. S. eds. (1991). *Cellular and Molecular Immunology.* W. B. Sauders Company.

Abremski, K., Hoess, R. and Sternberg, N. (1983). Studies on the properties of P1 site-specific recombination: Evidence for topologically unlinked products following recombination. *Cell*, **32**, 1301-1311.

Adams, A., Gottscling, D. and Kaiser, C. (1996). *Methods in Yeast Genetics- A laboratory course manual.* Cold Spring Harbor Laboratory Press.

Adams, B. G. (1972). Induction of galactokinase in *Saccharomyces cerevisiae:* Kinetics of induction and glucose effects. *J. Bacteriol.*, **111**, 308-315.

Alfthan, K., Takkinen, K., Sizmann, D., Söderlund, H. and Teeri, T. T. (1995). Properties of a single-chain antibody containing different linker peptides. *Protein Engng*, **8**, 725-731.

Alzari, P. M., Lascombe, M.-B. and Poljak, R. J. (1988). Three-dimensional structure of antibodies. *Ann. Rev. Immunol*, **6**, 555-580.

Amit, A. G., Mariuzza, R. A., Phillips, S. E. V. and Poljak, R. J. (1986) Three-dimensional structure of an antigen-antibody complex at 2.8 Å resolution. *Science*, **233**, 747-753.

Argos, P. (1990). An investigation of oligopeptides linking domains in protein tertiary structures and possible candidates for general gene fusions. *J. Mol. Biol.*, **211**, 943-958.

Austin, S., Ziese, M. and Sternberg, N. (1981). A novel role for site-specific recombination in maintence of bacterial replicons. *Cell*, **25**, 729-736.

Bagshawe, K. D. (1989). The first Bagshawe lecture - Towards generating cytotoxic agents at cancer sites. *Br. J. Cancer*, **60**, 275-281.

Barbas, C.F., III, Hu, D., Dunlop, N., Sawyer, L., Cababa, D., Hendry, R. M., Nara, P. L. and Burton, D. R. (1994). *In vitro* evolution of a neutralizing human antibody to human immunodeficiency virus type 1 to enhance affinity and broaden strain cross-reactivity. *Proc. Natl. Acad. Sci. USA.*, **91**, 3809-3813.

Bardwell, L., Cook, J.G., Inouge, C.J. and Thorner, J. (1994). Signal propagation and regulation in the mating pheromone response pathway of the yeast *Saccharomyces cerevisiae. Developmental biology,* **166**, 363-379.

Barnes, G. and Rine, J. (1985). Regulated expression of endonuclease *Eco* RI in *Saccharomyces cerevisiae*: Nuclear entry and biological consequences. *Proc. Natl. Acad. Sci.* USA, **82**, 1354-1358.

Beggs, J. D. (1978). Transformation of yeast by a replicating hybrid plasmid. *Nature*, **275**, 104-109.

Berek, C., Berger, A. and Apel, M. (1991). Maturation of the immune response in germinal centres. *Cell*, **67**, 1121-1129.

Berek, C. and Milstein, C. (1987). Mutation drift and repertoire shift in the maturation of the immune response. *Immunol. Rev.*, **93**, 23-41.

Berek, C. and Milstein, C. (1988). The dynamic nature of the antibody repertoire. *Immunol. Rev.*, **105**, 5-28.

Better, M., Chang, C.P., Robinson, R. R. and Horwitz, A.H. (1988). *Escherichia coli* secretion of an active chimeric antibody fragment. *Science*, **240**, 1041-1043.

Betz, A. G., Rada, C., Pannell, R., Milstein, C. and Neuberger, M. S. (1993). Passenger transgenes reveal intrinsic specificity of the antibody hypermutation mechanism: Clustering, polarity, and specific hot spots. *Proc. Natl. Acad. Sci. USA,* **90**, 2385-2388.

Betz, R., Duntze, W. and Manney, T.R. (1978). Mating-factor-mediated sexual agglutination in *Saccharomyces cerevisiae. FEMS Microbiol. Lett.*, **4**, 107-110.

Betz, R., Mackay, V. L. and Duntze, W. (1977). a-factor from *Saccharomyces cerevisiae*: Partial characterization of a mating hormone produced by cells of mating type **a**. *J. Bacteriol.,* **132**, 462-472.

Bhat, T. N., Bentley, G. A., Boulot, G., Greene, M. I., Tello, D., Dall'Acqua, W., Souchon, H., Schwarz, F. P., Mariuzza, R. A. and Poljak, R.J. (1994). Bound water molecules and conformational stabilization help mediate an antigen-antibody association. *Proc. Natl. Acad. Sci. USA.*, **91**, 1089-1093.

Bhat, T. N., Bentley, G. A., Fischmann, T. O., Boulot, G. and Poljak, R. J. (1990). Small rearrangements in structures of Fv and Fab fragments of antibody D1.3 on antigen binding. *Nature*, **347**, 483-485.

Bird, R. E., Hardman, K. D., Jacobson, J. W., Johnson, S., Kaufman, B. M., Lee, S.-M., Lee, T., Pope, S. H., Riordan, G. S. and Whitlow, M. (1988). Single-chain antigen-binding proteins. *Science,* **242**, 423-426.

Bird, R. E. and Walker, B. W. (1991). Single-chain antibody variable regions. *Tibtech,* **9**, 132-136.

Blake, C. C. F., Koenig, D. F., Mair, G. A., North, A. C. T., Phillips, D. C. and Sarma, V. R. (1965). Structure of hen egg-white lysozyme. A three-dimensional Fourier synthesis at 2 Å resolution. *Nature (London)*, **206**, 757-761.

Boder, E. T. and Wittrup, K. D. (1997). Yeast surface display for screening combinatorial polypeptide libraries. *Nature Biotechnol.,* **15**, 553-557.

Borrebaeck, C. A. K. ed. (1995). *Antibody Engineering* - 2nd Ed. Oxford University Press.

Boss, M. A., Kenten, J. H., Wood, C. R. and Emtage, J. S. (1984). Assembly of functional antibodies from immunoglobulin heavy and light chains synthesised in *E. coli. Nucleic Acids Res.,* **12**, 3791-3795.

Bowdish, K., Tang, Y., Hicks, J. B. and Hilvert, D. (1991). Yeast expression of a catalytic antibody with chorismate mutase activity. *J. Biol. Chem.,* **226**, 11901-11908.

Broach, J. R. (1982). The yeast plasmid 2μ circle. *Cell*, **28**, 203-204.

Cabilly, S., Riggs, A. D., Pande, H., Shively, J. E., Holmes, W. E., Rey, M., Perry, L.J., Wetzel, R. and Heyneker, H. L. (1984). Generation of antibody activity from immunoglobulin polypeptide chains produced in *E. coli. Proc. Natl. Acad. Sci. USA,* **81**, 3273-3277.

Camacho, S. A., Kosco-Vilbois, M. H. and Berek, C. (1998). The dynamic structure of the germinal center. *Immunol. Today,* **19**, 511-514.

Cappellaro, C., Baldermann, C., Rachel, R. and Tanner, W. (1994). Mating-type-specific cell-cell recognition of *Saccharomyces cerevisiae* α-agglutinin. *J. EMBO,* **13**, 4737-4744.

Cappellaro, C., Hauser, K., Mrsa, V., Watzele, M., Watzele, G., Gruber, C. and Tanner, W. (1991). *Saccharomyces cerevisiae* a- and α-agglutinin: characterisation of their molecular interaction. *J. EMBO,* **10**, 4081-4088.

Carlson, J. R. (1988). A new means of inducibly inactivating a cellular protein. *Mol. Cell. Biol.,* **88**, 2638-2646.

Castañón, M. J., Spevak, W., Adolf, G. R., Chlebowicz-Sledziewska, E. and Sledzieski, A. (1988). Cloning of human lysozyme gene and expression in the yeast *Saccharomyces cerevisiae*. *Gene*, **66**, 223-234.

Chen, M.-H., Shen, Z.-M., Bodin, S., Kahn, P. C. and Lipke, P. N. (1995). Structure of *Saccharomyces cerevisiae* α-agglutinin. *J. Biol. Chem.,* **270**, 26168-26177.

Chester, K. A. and Hawkins, R. E. (1995). Clinical issues in antibody design. *Tibtech,* **13**, 294-300.

Cho, B. K., Kieke, M. C., Boder, E. T., Wittrup, K. D. and Kranz, D. M. (1998). A yeast surface display system for the discovery of ligands that trigger cell activation. *J. Immunol. Methods.,* **220**, 179-188.

Chothia, C. and Lesk, A. M. (1987). Canonical structures of the hypervariable regions of immunoglobulins, *J. Mol. Biol.,* **196**, 901-917.

Chothia, C., Lesk, A. M., Tramontano, A., Levitt, M., Smith-Gill, S. J., Air, G., Sheriff, S., Padlan, E. A., Davies, D., Tulip, W. R., Colman, P. M., Spinelli, S., Alzari, P. M. and Poljak, R. J. (1989). Conformations of immunoglobulin hypervariable regions, *Nature*, **342**, 877-883.

Chothia, C., Lesk, A. M., Gherardi, E., Tomlinson, I. M., Walter, G., Marks, J. D., Llewelyn, M. B. and Winter, G. (1992). Structural repertoire of the human VH segments. *J. Mol. Biol.,* **227**, 799-817.

Clackson, T., Hoogenboom, H. R., Griffiths, A. D. and Winter, G. (1991). Making antibody fragments using phage display libraries. *Nature*, **352**, 624-628.

Cook, G.P. and Tomlinson, I.M. (1995). The human immunoglobulin VH repertoire. *Immunology Today*, **16**, 237-242.

Cox, J. P. L., Tomlinson, I. M. and Winter, G. (1994). A directory of human germ-line Vκ segments reveals a strong bias in their usage. *Eur. J. Immunol.*, **24**, 827-836.

Davies, D. R., Padlan, E. A. and Sheriff, S. (1990). Antibody-antigen complexes. *Ann. Rev. Biochem.*, **59**, 439-473.

de Nobel, H., Lipke, P. N. and Kurjan, J. (1996). Identification of a ligand-binding site in an immunoglobulin fold domain of the *Saccharomyces cerevisiae* adhesion protein α-agglutinin. *Molecular Biology of the Cell*, **7**, 143-153.

Desplancq, D., King, D. J., Lawson, A. D. G. and Mountain, A. (1994). Multimerization behaviour of single chain Fv variants for the tumour binding antibody B72.3. *Protein Eng*, **7**, 1027-1033.

Doi, S., Tanabe, K., Watanabe, M., Yamaguchi, M. and Yoshimura, M. (1989). An α-specific gene, *SAG1* is required for sexual agglutination in *Saccharomyces cerevisiae. Curr. Genet.*, **15**, 393-398.

England, P., Nageotte, R., Renard, M., Page, A. -L. and Bedouelle, H. (1999). Functional characterization of the somatic hypermutation process leading to antibody D1.3, a high affinity antibody directed against lysozyme. *J. Immunol.*, **162**, 2129-2136.

EMBO practical course, Pomezia. (1991). Making antibodies in bacteria and on phage. Medical Research Council, Cambridge Centre for Protein Engineering.

Fischmann, T. O., Bentley, G. A., Bhat, T. N., Boulot, G., Mariuzza, R. A., Phillips, S. E. V., Tello, D. and Poljak, R. J. (1991). Crystallographic refinement of the three-dimensional structure of the Fab D1.3-lysozyme complex at 2.5Å resolution. *J. Biol. Chem.*, **266**, 12915-12920.

Foote, J. and Winter, G. (1992). Antibody framework residues affecting conformation of the hypervariable loops. *J. Mol. Biol.*, **224**, 487-499.

Foury, F., Roganti, T., Lecrenier, N. and Purnelle, B. (1998). The complete sequence of the mitochondrial genome of *Saccharomyces cerevisiae. FEBS Letts.*, **440**, 325-331.

Francisco, J. A., Campbell, R., Iverson, B. L. and Georgiou, G. (1993). Production and fluorescence-activated cell sorting of *E. coli* expressing a functional antibody fragment on the external surface. *Proc. Natl. Acad. Sci. USA*, **90**, 10444-10448.

Frippiat, J. -P., Williams, S. C., Tomlinson, I.M., Cook, G. P., Cherif, D., Le Paslier, D., Collins, J. E., Dunham, I., Winter, G. and Lefranc, M. -P. (1995). Organization of the human immunoglobulin lambda light chain locus on chromosome 22q11.2. *Human Molecular Genetics,* **4**, 983-991.

Fuchs, P., Breitling, F., Dübel, S., Seehaus, T. and Little, M. (1991). Targeting recombinant antibodies to the surface of *E. coli*: Fusion to a peptidoglycan associated lipoprotein. *Biotech,* **9**, 1369-1372.

Gerbaud, C., Fournier, P., Blanc, H., Aigle, M., Heslot, H. and Guerineau, M. (1979). High frequency of yeast transformation by plasmids carrying part or entire 2μm yeast plasmid. *Gene*, **5**, 233-253.

Ghetie, M. -A. and Vitetta, E. S. (1994). Recent developments in immunotoxin therapy. *Curr. Opin. Immunol.,* **6**, 707-714.

Glockshuber, R., Malia, M., Pfitzinger, I. and Plückthun, A. (1990). A comparsion of strategies to stablize immunoglobulin Fv-fragments. *Biochemistry,* **29**, 1362-1367.

Goffeau, A., Aert, R., Agostini-Carbone, M. L., Ahmed, A., Aigle, M., Alberghina, L., Albermann, K. *et al.*, (1997). The yeast genome directory. *Nature,* **387** (Suppl.), 5-105.

Gonzalez-Fernandez, A. and Milstein, C. (1998). Low antigen dose favours selection of somatic mutants with hallmarks of antibody affinity maturation. *Immunology*, **93**, 149-153.

Green, N. S., Lin, M. M. and Scharff, M. D. (1998). Somatic hypermutation of antibody genes: a hot spot warms up. *Bioessays*, **20**, 227-234.

Griffiths, A. D., Williams, S. C., Hartley, O., Tomlinson, I. M., Waterhouse, P., Crosby, W. L., Kontermann, R. E., Jones, P. T., Low, N. M., Allison, T. J., Prospero, T. D., Hoogenboom, H. R., Nissim, A., Harrison, J. L., Zaccolo, M., Gherardi, E. and Winter, G. (1994). Isolation of high affinity human antibodies directly form large synthetic repertoires. *J. EMBO,* **13**, 3245-3260.

Güldener, U., Heck, S., Fiedler, T.,Beinhauer, J. and Hegemann, J. H. (1996). A new efficient gene disruption cassette for repeated use in budding yeast. *Nucleic Acids Res.*, **24**, 2519-2524.

Guo, F., Gopaul, D. N. and Van Duyne, G. D. (1997). Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse. *Nature*, **389**, 40-46.

Hadfield, C., Jordan, B. E., Mount, R. C., Pretorius, G. H. J. and Burak, E. (1990). G418-resistance as a dominant marker and reporter for gene expression in *Saccharomyces cerevisiae*. *Curr. Genet.*,**18**, 303-313.

Hamada, K., Terashima, H., Arisawa, M. and Kitada, K. (1998). Amino acid sequence requirement for efficient incorporation of glycosylphosphatidylinositol-associated proteins into the cell wall of *Saccharomyces cerevisiae*. *J.Biol. Chem.*, **273**, 26946-26953.

Hamers-Casterman, C., Atarhouch, T., Muyldermans, S., Robinson, G., Hamers, C., Songa, E. B., Bendahman, N. and Hamers, R. (1993). Naturally occurring antibodies devoid of light chains, *Nature*, **363**, 446-448.

Hammond, C. and Helenius, A. (1995). Quality control in the secretory pathway. *Curr. Opin. Cell. Biol.*, **7**, 523-529.

Han, S., Zheng, B., Dal Porto, J. and Kelsoe, G. (1995). In situ studies of the primary immune response to (4-hydroxy-3-nitrophenyl) acetyl IV. Affinity-dependent antigen-driven B cell apoptosis in germinal centres as a mechanism for maintaining self-tolerance. *J. Exp. Med.*, **182**, 1635-1644.

Harper, M., Lema, F., Boulot, G. and Poljak, R. J. (1987). Antigen specificity and cross reactivity of monoclonal anti-lysozyme antibodies. *Mol. Immunol.*, **24**, 97-108.

Hauser, K. and Tanner, W. (1989). Purification of the inducible α-agglutinin of *S. cerevisiae* and molecular cloning of the gene. *FEBS Letters*, **255**, 290-294.

Hawkins, R. E., Russell, S. J., Baier, M. and Winter, G. (1993). The contribution of contact and non-contact residues of antibody in the affinity of binding to antigen: The interaction of mutant D1.3 antibodies with lysozyme. *J. Mol. Biol.*, **234**, 958-964.

Hawkins, R. E., Russell, S. J. and Winter, G. (1992) Selection of phage antibodies by binding affinity. *J. Mol. Biol.*, **226**, 889-896.

Herron, J. N., He, X. M., Ballard,D. W., Blier, P.R., Pace, P. E., Bothwell, A. L. M., Voss, E. W. and Edmundson, A. B. (1991). An autoantibody to single-stranded DNA: Comparsion of the three-dimensional structure of the unliganded Fab and a deoxynucleotide-Fab complex. *Proteins: Structure, Function and Genetics*, **11**, 159-175.

Herskowitz, I. (1988). Life cycle of the budding yeast, *Saccharomyces cerevisiae*. *Microbiol. Rev.*, **52**, 536-553.

Hoess, R.H. and Abremski, K. (1985). Mechanism of strand cleavage and exchange in the Cre-lox site-specific recombination system. *J. Mol. Biol.*, **181**, 351-362.

Hoess, R. H., Wierzbicki, A. and Abremski, K. (1986). The role of the *loxP* spacer region in P1 site-specfic recombination. *Nucleic Acids Res.*, **14**, 2287-2300.

Holliger, P., Prospero, T. and Winter, G. (1993). 'Diabodies': Small bivalent and bispecific antibody fragments. *Proc. Natl. Acad. Sci. USA*, **90**, 6444-6448.

Holliger, P., Wing, M., Pound, J. D., Bohlen, H. and Winter, G. (1997). Retargeting serum immunoglobulin with bispecific diabodies. *Nature Biotech.*, **15**, 632-636.

Holmes, M. A. and Foote, J. (1997). Structural consequences of humanizing an antibody. *J. Immunol.*, **158**, 2192-2201.

Hoogenboom, H.R. (1997). Designing and optimizing library selection strategies for generating high affinity antibodies. *Tibtech.*, **15**, 62-70.

Hoogenboom, H. R., Griffiths, A. D., Johnson, K. S., Chiswell, D. J., Hudson, P. and Winter, G. (1991). Multi-subunit proteins on the surface of filamentous phage: methodologies for displaying antibody (Fab) heavy and light chains. *Nucleic. Acids. Res.*, **19**, 4133-4137.

Hoogenboom, H. R. and Winter, G. (1992). By-passing immunization: Human antibodies from synthetic repertoires of germline VH gene segments rearranged *in vivo*. *J. Mol. Biol.*, **227**, 381-388.

Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K. and Pease, L. R. (1989). Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene*, **77**, 61-68.

Horwitz, A. H., Chang, C.P., Better, M., Hellstrom, K.E. and Robinson, R. R. (1988). Secretion of functional antibody and Fab fragments from yeast cells. *Proc. Natl. Acad. Sci. USA*, **85**, 8678-8682.

Huse, W. D., Sastry, L., Iverson, S. A., Kang, A. S., Alting-Mees, M., Burton, D. R., Benkovic, S. J. and Lerner, R. A. (1989). Generation of a large combinatorial library of the immunoglobulin repertoire in phage lambda. *Science*, **246**, 1275-1281.

Huston, J. S., Levinson, D., Mudgett-Hunter, M., Tai, M. -S., Novotný, J., Margolies, M. N., Ridge, R. J., Bruccoleri, R. E., Haber, E., Crea, R. and Oppermann, H. (1988). Protein engineering of antibody binding sites: Recovery of specific activity in an anti-digoxin single-chain Fv analogue in *E. coli. Proc. Natl. Acad. Sci. USA.*, **85**, 5879-5883.

Huston, J. S., McCartney, J., Tai, M. S., Mottola-Hartshom, C., Jin, D., Warren, F., Keck, P. and Opperman, H. (1993). Medical applications of single chain antibodies. *Intern. Rev. Immunol.*, **10**, 195-217.

Huston, J. S., Mudgett-Hunter, M., Tai, M.-S., McCartney, J., Warren, F., Haber, E. and Oppermann, H. (1991). Protein engineering of single-chain Fv analogs and fusion proteins. *Methods in Enzymology*, **203**, 46-87.

Jackson, C. L. and Hartwell, L. H. (1990). Courtship in *S.cerevisiae*: Both cell types choose mating partners by responding to the strongest pheromone signal. *Cell*, **63**, 1039-1051.

Jacob, J., Kelsoe, G., Rajewsky, K. and Weiss, U. (1991). Intraclonal generation of antibody mutants in germinal centres. *Nature*, **354**, 389-392.

James, K. and Bell, G. T. (1987). Human monoclonal antibody production: Current status and future prospects. *J. Immunol. Methods*, **100**, 5-40.

Jigami, Y., Muraki, M., Harada, N. and Tanaka, H. (1986). Expression of synthetic human-lysozyme gene in *Saccharomyces cerevisiae:* Use of a synthetic chicken-lysozyme signal sequence for secretion and processing. *Gene*, **43**, 273-279.

Johnson, S. and Bird, R. E. (1991). Construction of single-chain Fv derivatives of monoclonal antibodies and their production in *E. coli. Methods Enzymol.*, **203**, 88-98.

Johnston, J. R. ed. (1994). *Molecular Genetics of Yeast - A practical approach.* Oxford University Press.

Johnston, M. (1987). A model fungal gene regulatory mechanism: the *GAL* genes of *Saccharomyces cerevisiae. Microbiol. Rev.*, **51**, 458-476.

Jones, E. W. in ' Guide to yeast genetics and molecular biology' Chapter 31. (1991). Guthrie, C. and Fink, G. R. eds. Academic Press.

Jones, P. T., Dear, P. H., Foote, J., Neuberger, M. S. and Winter, G. (1986). Replacing the complementarity-determining regions in a human antibody with those from a mouse. *Nature,* **321**, 522-525.

Jost, C. R., Kurucz, I., Jacobus, C. M., Titus, J. A., George, A. J. T. and Segal, D. M. (1994). Mammalian expression and secretion of functional single-chain Fv molecules. *J. Biol. Chem.,* **269**, 26267-26273.

Jung, A., Sippel, A. E., Grez, M. and Schütz, G. (1980). Exons encode functional and structural units of chicken lysozyme. *Proc. Natl. Acad. Sci. USA*, **77**, 5759-5763.

Kabat, E. A., Wu, T. T., Perry, H. M., Gottesmann, K. S. and Foeller, C. (1991) *Sequences of Proteins of Immunological Interest.* 5th edn. US Dept of Health and Human Services, Bethesda, USA.

Kang, A. S., Barbas, C. F., Janda, K. D., Benkovic, S. J. and Lerner, R. A. (1991). Linkage of recognition and replication functions by assembling combinational antibody Fab libraries along phage surfaces. *Proc. Natl. Acad. Sci. USA.,* **88**, 4363-4366.

Kato, A., Takasaki, H. and Ban, M. (1994a). Polymannosylation to asparagine-19 in hen egg-white lysozyme in yeast. *FEBS letters*, **355**, 76-80.

Kato, A., Tanimoto, S., Muraki, Y., Oda, Y., Inoue, Y. and Kobayashi, K. (1994b). Relationships between conformational stabilities and surface functional properties of mutant hen egg-white lysozyme construct by genetic engineering. *J. Agric. Food Chem.*, **42**, 227-230.

Kelsoe, G. (1996). Life and death in germinal centres. *Immunity*, **4**, 107-111.

Kepler, T. B. and Perelson, A. S. (1993). Cyclic re-entry of germinal centre B-cells and efficiency of affinity maturation. *Immunol. Today*, **14**, 412-415.

Kieke, M. C., Cho, B. K., Boder, E. T., Kranz, D. M. and Wittrup, K. D. (1997). Isolation of anti-T cell receptor scFv mutants by yeast surface display. *Protein Eng.*, **10**, 1303-1310.

King, D. J., Turner, A., Farnsworth, A. P. H., Adair, J. R., Owens, R. J. *et al.* (1994). Improved tumor targeting with chemically cross-linked recombinant antibody fragments. *Cancer Res.*, **54**, 6176-6185.

Kipriyanov, S. M., Moldenhauer, G., Martin, A. C. R., Kupriyanova, O. A. and Little, M. (1997). Two amino acid mutations in an anti-human CD3 single chain Fv antibody fragment that affect the yield of bacterial secretion but not the affinity. *Protein Eng.*, **10**, 445-453.

Klis, F. M. (1994). Review: Cell wall assembly in yeast. *Yeast*, **10**, 851-869.

Klis, F.M., Schreuder, M.P., Toschka, H.Y. and Verrips, C.T. (1994). *WO patent* 94/01567

Knappik, A. and Plückthun, A. (1995). Engineered turns of a recombinant antibody improve its *in vivo* folding. *Protein Eng.*, **8**, 81-89.

Kocks, C. and Rajewsky, K. (1989). Stable expression and somatic hypermutation of antibody V regions in B-cell development pathways. *Ann. Rev. Immunol.*, **7**, 537-559.

Köhler, G. and Milstein, C. (1975). Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature*, **256**, 495-497.

Kontermann, R. E., Wing, M. G. and Winter, G. (1997). Complement recruitment using bispecific diabodies. *Nature Biotech.*, **15**, 629-631.

Kosco, M. H., Szakal, A.K. and Tew, J.G. (1988). *In vivo* obtained antigen presented by germinal centre B cells to T cells *in vitro*. *J. Immunol*, **140**, 354-360.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bactriophage T4. *Nature*, **227**, 680-685.

Lakso, M., Sauer, B., Mosinger, Jr., B., Lee, E. J., Manning, R. W., Yu, S.-H., Mulder, K. L. and Westphal, H. (1992). Targeted oncogene activation by site-specific recombination in transgenic mice. *Proc. Natl. Acad. Sci. USA*, **8**, 6232-6236.

Längle-Rouault, F. and Jacobs, E. (1995). A method for performing precise alternations in the yeast genome using a recyclable selectable marker. *Nucleic Acids Res.* **23**, 3079-3081.

Lanzavecchia, A. (1985). Antigen-specific interaction between T and B cells. *Nature*, **314**, 537-539.

Lesk, A. M. and Chothia, C. (1982). Evolution of proteins formed by β-sheets. II. The core of the immunoglobulin domains. *J. Mol. Biol.*,**160**, 325-342.

Li, Y., Owen, M. R .L., Cockburn, W., Kurragai, I. and Whitelam, G. C. (1996). Study of antibody-antigen interaction through site-directed mutagenesis of the VH region of a hybrid phage-antibody fragment. *Protein. Eng.*, **9**, 1211-1217.

Lindhout, E., Koopman, G., Pals, S. T. and de Groot, C. (1997). Triple check for antigen specificity of B cells during germinal centre reactions. *Immunol. Today*, **18**, 573-577.

Lindhout, E., Lakeman, A. and de Groot, C. (1985). Follicular dendritic cells inhibit apoptosis in human B-lymphocytes by a rapid and irreversible blockade of preexisting endonuclease. *J. Exp. Med.*, **181**, 1985-1995.

Lipke, P. N. and Kurjan, J. (1992). Sexual agglutination in budding yeasts: Structure, function, and regulation of adhesion glycoproteins. *Microbiol. Rev.*, **56**, 180-194.

Lipke, P. N., Terrance, K. and Wu, Y.-S. (1987). Interaction of α-agglutinin with *Saccharomyces cerevisiae* **a**-cells. *J. Bacteriol.* **169**, 483-488.

Lipke, P. N., Wojciechowicz, D. and Kurjan, J. (1989). AGα1 is the structural gene for the *Saccharomyces cerevisiae* α-agglutinin, a cell surface glycoprotein involved in cell-cell interactions during mating. *Mol. Cell. Biol.* **9**, 3155-3165.

Liu, Y.-J. and Arpin, C. (1997). Germinal centre development. *Immunol.Rev.*, **156**, 111-126.

Liu, Y.-J., Joshua, D. E., Williams, G. T., Smith, C. A., Gordon, J. and MacLennan, I.C.M. (1989). Mechanism of antigen-driven selection in germinal centres. *Nature*. **342**, 929-931.

Low, N. M., Holliger, P. and Winter, G. (1996). Mimicking somatic hypermutation: Affinity maturation of antibodies displayed on bacteriophage using a bacterial mutator strain. *J. Mol. Biol.*, **260**, 359-368.

Lu, C.-F., Kurjan, J. and Lipke, P. N. (1994). A pathway for cell wall anchorage of *Saccharomyces cerevisiae* α-agglutinin. *Mol. Cell. Biol.*, **14**, 4825-4833.

Lu, C.-F., Montijn, R. C., Brown, J. L., Klis, F., Kurjan, J., Bussey, H. and Lipke, P. N. (1995). Glycosyl phosphatidylinostol-dependent cross-linking of α-agglutinin and β-1,6-glucan in the *Saccharomyces cerevisiae* cell wall. *J. Cell. Biol.*, **128**, 333-340.

Mackay, V. L., Welch, S. K., Insley, M. Y., Manney, T. R., Holly, J., Saari, G. C., and Parher, M. L. (1988). The *Saccharomyces cerevisiae BAR1* gene encodes an exported protein with homology to pepsin. *Proc. Natl. Acad. Sci. USA.*, **85**, 55-59.

Maclennan, I.C.M. (1994). Germinal centres. *Ann. Rev. Immunol.*, **12**,117-139.

Malmborg, A-C., Söderlind, E., Frost, L. and Borrebaeck, C. A. K. (1997). Selective phage infection mediated by epitope expression on F pilus. *J. Mol. Biol.*, **273**, 544-551.

Marks, J. D., Griffiths, A. D., Malmqvist, M., Clackson, T. P., Bye, J. M., and Winter, G. (1992). By-passing immunization: Building high affinity human antibodies by chain shuffling. *Bio/tech.*, **10**, 779-783.

Marks, J. D., Hoogenboom, H.R., Bonnert, T.P., McCafferty, J., Griffiths, A. D. and Winter, G. (1991). By-passing immunization: Human antibodies from V-gene libraries displayed on phage. *J. Mol. Biol.*, **222**, 581-597.

McCafferty, J., Griffiths, A. D., Winter, G. and Chiswell, D. J. (1990). Phage antibodies: filamentous phage displaying antibody variable domains. *Nature*, **348**, 552-554.

McPherson, M. J., Quirke, P. and Taylor, G. R. eds. (1992). *PCR - A practical approach.* Oxford University Press.

Michaelis, S. and Herskowitz, I. (1988). The a-factor pheromone of *Saccharomyces cerevisiae* is essential for mating. *Mol. Cell. Biol.*, **8**, 1309-1318.

Morea, V., Tramontano, A., Rustici, M., Chothia, C. and Lesk, A. M. (1998). Conformations of the third hypervariable region of the VH domain of immunoglobulins. *J. Mol. Biol.*, **275**, 269-294.

Morrison, S. L., Johnson, M. J., Herzenberg, L. A. and Oi, V. T. (1984). Chimaeric human antibody molecules: mouse antigen-binding domains. *Proc. Natl. Acad. Sci. USA.*, **82**, 6851-6855.

Mörsky, P. (1983). Turbidimetric determination of lysozyme with *Micrococcus lysodeikticus* cells: Reexamination of reaction conditions. *Anal. Biochem.*, **128**, 77-85.

Mortimer, R. K. and Johnston, J. R. (1986). Genealogy of principal strains of the yeast genetic stock centre. *Genetics*, **113**, 35-43.

Murai, T., Ueda, M., Atomi, H., Shibasaki, Y., Kamasawa, N., Osumi, M., Kawaguchi, T., Arai, M. and Tanaka, A. (1997a). Genetic immobilization of cellulase on the cell surface of *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.*, **48**, 499-503.

Murai, T., Ueda, M., Yamamura, M., Atomi, H., Shibasaki, Y., Kamasawa, N., Osumi, M., Amachi, T. and Tanaka, A. (1997b). Construction of a starch-utilizing yeast by cell surface engineering. *Appl. Environ. Microbiol.*, **63**, 1362-1366.

Muraki, M., Jigami, Y., Tanaka, H., Kishimoto, F., Agui, H., Ogino, S. and Nakasato, S. (1985). Expression of synthetic human lysozyme gene in *Escherichia coli*. *Agric. Biol. Chem.*, **49**, 2829-2831.

Muyldermans, S. and Lauwereys, M. (1999). Unique single-domain antigen binding fragments derived from naturally occuring camel heavy-chain antibodies. *Journal of Molecular Recognition*, **12**, 131-140.

Nagai, K. and Thogersen, H. C. (1984). Generation of β-globin by sequence-specific proteolysis of a hybrid protein produced in *Escherichia coli*. *Nature*, **309**, 810-812.

Neri, D., Carnemolla, B., Nissim, A., Leprini, A., Querzè, G., Balza, E., Pini, A., Tarli, L., Halin, C., Nerl, P., Zardi, L. and Winter, G. (1997). Targeting by affinity-matured recombinant antibody fragments of an angiogenesis associated fibronectin isoform. *Nature Biotech.*, **15**, 1271-1275.

Neuberger, M. S. and Milstein, C. (1995). Somatic hypermutation. *Curr. Opin. Immunol.*, **7**, 248-254.

Nieba, L., Honegger, A., Krebber, C. and Plückthun, A. (1997). Disrupting the hydrophobic patches at the antibody variable/constant domain interface: Improved *in vivo* folding and physical characteristic of an engineered scFv fragment. *Protein Eng.*, **10**, 435-444.

Oberto, J. and Davison, J. (1985). Expression of chicken egg-white lysozyme by *Saccharomyces cerevisiae*. *Gene*, **40**, 57-65.

Orlandi, R., Güssow, D. H., Jones, P. T. and Winter, G. (1989). Cloning immunoglobulin variable domains for expression by the polymerase chain reaction, *Proc. Natl. Acad. Sci. USA*, **86**, 3833-3837.

Oriean, P., Kuranda, M. J. and Albright, C. F. in ' Guide to yeast genetics and molecular biology '. (1991). Chapter 47. Guthrie, C and Fink, G. R. eds. Academic Press.

Oprea, M. and Perelson, A. S. (1997). Somatic mutation leads to efficient affinity maturation when centrocytes recycle back to centroblasts. *J. Immunol.*, **158**, 5155-5162.

Osserman, E. F. and Lawlor, D. P. (1966). Serum and urinary lysozyme (muramidase) in monocytic and monomyelocytic leukemia. *J. Exp. Med.* **124**, 921-951.

Owen. M., Gandecha, A., Cockburn, B. and Whitelam, G. (1992). Synthesis of a functional anti-phytochrome single-chain Fv protein in transgenic tobacco. *Bio/tech*, **10**, 790-794.

Pascual, V., Liu, Y. J., Magalski, A., deBouteiller, O., Banchereau, J. and Capra, J. D. (1994). Analysis of somatic mutation in five B cell subsets of human tonsil. *J. Exp. Med.*, **180**, 329-339.

Perelson, A. S. and Oster, G. F. (1979). Theoretical studies of clonal selection: Minimal antibody repertoire size and reliability of self and non-self discrimination. *J. Theor. Biol.*, **81**, 645-670.

Peters, A. and Storb, U. (1996). Somatic hypermutation of immunoglobulin genes is linked to transcription initiation. *Immunity*, **4**, 57-65.

Plückthun, A. (1991). Antibody engineering: Advances from the use of *Escherichia coli* expression systems. *Bio/tech*, **9**, 545-551.

Poljak, R. J., Amzel, L. M., Avey, H. P., Chen, B. L., Phizackerly, R. P. and Saul, F. (1973). Three-dimensional structure of the Fab' fragment of the human immunoglobulin at 2.8Å resolution. *Proc. Natl. Acad. Sci. USA.*, **70**, 3305-3310.

Pringle, J. R., Adams, A. E. M., Drubin, D. G. and Haarer, B. K. in ' Guide to yeast genetics and molecular biology'. (1991). Chapter 40. Guthrie, C. and Fink, G. R. eds. Academic Press.

Proteau, G., Gelaznikas, R. and Merante, F. (1990). The isolation of biologically active mating pheromone, a-factor, from the yeast, *Saccharomyces cerevisiae*. *Biochem. Biophys. Res. Commun.*, **170**, 182-186.

Qin, M., Bayley, C., Stockton, T. and Ow, D.W. (1994) Cre recombinase-mediated site-specific recombination between plant chromosomes. *Proc. Natl. Acad. Sci. USA*, **91**, 1706-1710.

Queen, C., Schneider, W. P., Selick, H. E., Payne, P. W., Landolfi, N. F., Duncan, J. F., Avdalovic, N. M., Levitt, M., Junghans, R. P. and Waldmann, T. A. (1989). A humanized antibody that binds to the interleukin-2 receptor. *Proc. Natl. Acad. Sci. USA,* **86**, 10029-10033.

Roguska, M. A., Pedersen, J. T., Henry, A. H., Searle, S. M. J., Roja, C. M., Avery, B., Hoffee, M., Cook, S., Lambert, J. M., Blättler, W.A., Rees, A. R. and Guild, B. C. (1996). A comparsion of two murine monoclonal antibodies humanized by CDR-grafting and variable domain resurfacing. *Protein. Eng,* **9**, 895-904.

Roitt, I., Brostoff, J. and Male, D. eds. (1986). *Immunology.* Gower Medical Publishing, London.

Romanos, M. A., Scorer, C. A. and Clare, J. J. (1992). Foreign gene expression in yeast: a review. *Yeast*, **8**, 423-488.

Ross-Macdonald, P., Sheehan, A., Roeder, G. S. and Snyder, M. (1997). A multipurpose transposon system for analysing protein production, localization, and function in *Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. USA.,* **94**, 190-195.

Rothstein, R. J. (1983). One-step gene disruption in yeast. *Methods Enzymol.*, **101**, 202-211.

Roy, A., Lu, C. F., Marykwas, D. L., Lipke, P. N. and Kurjan, J. (1991). The *AGA1* product is involved in cell surface attachment of the *Saccharomyces cerevisiae* cell adhesion glycoprotein a-agglutinin. *Mol. Cell. Biol.*, **11**, 4196-4206.

Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G.T., Erlich, H. A., and Arnheim, N. (1985). Enzymatic amplification of β-globulin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science,* **230**, 1350-1354.

Sambrook, J., Fritsch, E. F. and Maniatis, T. eds. (1989). *Molecular cloning. A laboratory manual.* Cold Spring Harbor Laboratory Press.

Sander, P., Grünewald, S., Bach, M., Haase, W., Reiländer, H. and Michel, H. (1994). Heterologous expression of the human D_{2S} dopamine receptor in protease-deficient *Saccharomyces cerevisiae* strains. *Eur. J. Biochem.,* **226**, 697-705.

Sanger, F., Nicken, S. and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.,* **74**, 5463-5467.

Sauer, B. (1987). Functional expression of the Cre-*lox* site-specific recombination system in the yeast *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **7**, 2087-2096.

Sauer, B. (1992). Identification of cryptic *lox* sites in the yeast genome by selection for Cre-mediated chromosome translocations that confer multiple drug resistance. *J. Mol. Biol.*, **223**, 911-928.

Sauer, B. (1993). Manipulation of transgenes by site-specific recombination: Use of Cre recombinase. *Methods Enzymol.* **225**, 890-900.

Schäble, K. F. and Zachau, H.-G. (1993). The variable genes of the human immunoglobulin κ locus. *Biol. Chem. Hoppe-Seyler.*, **374**, 1001-1022.

Scherer, S. and Davis, R. W. (1979). Replacement of chromosome segments with altered DNA sequences constructed *in vitro*. *Proc. Natl. Acad. Sci.*, **76**, 4951-4955.

Schier, R., Balint, R. F., McCall, A., Apell, G., Larrick, J. W. and Marks, J. D. (1996). Identification of functional and structural amino-acid residues by parsimonious mutagenesis. *Gene,* **169**, 147-155.

Schiestl, R. H., Manivasakam, P., Woods, R. A. and Gietz, R. D. (1993). Introducing DNA into yeast by transformation. *Methods : A companion to Methods in Enzymology*, **5**, 79-85.

Schlom, J., Eggensperger, D., Colcher, D., Molinolo, A., Houchens, D., Miller, L. S., Hinkle, G. and Siler, K. (1992). Therapeutic advantage of high anticarcinoma radioimmunoconjugates. *Cancer Res.*, **52**, 1067-1072.

Schreuder, M. P., Brekelmans, S., van den Ende, H. and Klis, F. M. (1993). Targeting of a heterologous protein to the cell wall of *Saccharomyces cerevisiae*. *Yeast,* **9**, 399-409.

Schreuder, M. P., Deen, C., Boersma, W. J. A., Pouwels, P. H. and Klis, F. M. (1996). Yeast expressing hepatitis B virus surface antigen determinants on its surface: Implications for a possible oral vaccine. *Vaccine*, **14**, 383-388.

Sheriff, S. and Constantine, K. L. (1996). Redefining the minimal antigen-binding fragment. *Nature Struct. Biology,* **3**, 733-736.

Sherman, F. in 'Guide to yeast genetics and molecular biology' (1991). Chapter 1. Guthrie, C. and Fink, G. R. eds. Academic Press.

Shokat, K. M. and Goodnow, C. C. (1995). Antigen-induced B-cell death and elimination during germinal centre immune responses. *Nature*, **375**, 334-338.

Shusta, E. V., Raines, R. T., Plückthun, A. and Wittrup, K. D. (1998). Increasing the secretory capacity of *Saccharomyces cerevisiae* for production of single-chain antibody fragments. *Nat. Biotechnol.*, **16**, 773-777.

Sikorski, R. S. and Hieter, P. (1989). A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae. Genetics*, **122**, 19-27.

Singh, A., Chen, E. Y., Lugovoy, J. M., Chang, C. N., Hitzeman, R. E. and Seeburg, P. H. (1983). *Saccharomyces cerevisiae* contains two discrete genes coding for the α-factor pheromone. *Nucleic. Acids. Res.*,**11**, 4049-4063.

Skerra, A. and Plückthun, A. (1988). Assembly of a functional immunoglobulin Fv fragment in *E. coli. Science,* **240**, 1038-

Sophianopoulous, A. J., Rhodes, C. K., Holcomb, D. N. and Van Holde, K. E. (1962). Physical studies of lysozyme. *J. Biol. Chem.,* **237**, 1107-1112.

Sprague, G. F. in 'Guide to yeast genetics and molecular biology' (1991). Chapter 5. Guthrie, C. and Fink, G. R. eds. Academic press.

Stemmer, W. P. C. (1994). Rapid evolution of a protein *in vitro* by DNA shuffling. *Nature*, **370**, 389-391.

Sternberg, N. and Hamilton, D. (1981). Bacteriophage P1 Site-Specific recombination I. Recombination between *loxP* sites. *J. Mol. Biol.,* **150**, 467-486.

Sternberg, N., Hamilton, D. and Hoess, R. (1981). Bacteriophage P1 site-specific recombination. II. Recombination between *loxP* and the bacterial chromosome. *J. Mol. Biol.,* **150**, 487-507.

Sternberg, N., Sauer, B., Hoess, R. and Abremski, K. (1986). Bacteriophage P1 Cre gene and its regulatory region. Evidence for multiple promoters and for regulation by DNA methylation. *J. Mol. Biol.,* **187**, 197-212.

Szakal, A. K., Kosco, M. H. and Tew, J. G. (1988). A novel *in vivo* follicular dendritic cell-dependent iccosome-mediated mechanism for delivery of antigen to antigen-processing cells. *J. Immunol.*, **140**, 341-353.

Szakal, A. K., Kosco, M. H. and Tew, J. G.(1989). Microanatomy of lymphoid tissue during humoral immune responses: Structure Function Relationships. *Ann. Rev. Immunol*, **7**, 91-109.

Szakal, A. K. and Tew, J. G. (1992). Follicular dendritic cells: B-cell proliferation and maturation. *Cancer Res. (Suppl.)*, **52**, 5554-5556.

Tang, Y., Hicks, J. B. and Hilvert, D. (1991). *In vivo* catalysis of a metabolically essential reaction by an antibody. *Proc. Natl. Acad. Sci. USA*, **88**, 8764-8786.

Terrance, K., Heller, P., Wu, Y.-S. and Lipke, P. N. (1987). Identification of glycoprotein components of α-agglutinin, a cell adhesion protein from *Saccharomyces cerevisiae. J. Bacteriol.* **169**, 475-482.

Terrance, K. and Lipke, P. N. (1981). Sexual Agglutination in *Saccharomyces cerevisiae. J. Bacteriol.* **148**, 889-896.

Terrance, K. and Lipke, P. N. (1987). Pheromone induction of sexual agglutinability in *Saccharomyces cerevisiae* **a** cells. *J. Bacteriol.* ,**169**, 4811-4815.

Tew, J.G., Kosco, M. H. and Szakal, A.K. (1989). The alternative antigen pathway. *Immunol. Today*, **10**, 229-234.

Tew, J. G., Wu, J., Qin, D., Helm, S., Burton, G. F. and Szakal, A. K. (1997). Follicular dendritic cells and presentation of antigen and costimulatory signals to B cells. *Immunol. Rev.*, **156**, 39-52.

Thim, L., Hansen, M. T., Norris, K., Hoegh, I., Boel, E., Forstrom, J. Ammerer, G. and Fiil, N. P. (1986). Secretion and processing of insulin precusors in yeast. *Proc. Natl. Acad. Sci. USA*, **83**, 6766-6770.

Tohoyama, H. and Yanagishima, N. (1985). The sexual agglutination substance is secreted through the yeast secretory pathway in *Saccharomyces cerevisiae. Mol. Gen. Genet.*, **201**, 446-449.

Tomlinson, I. M., Walter, G., Jones, P. T., Dear, P. H., Sonnhammer, E. L. L. and Winter, G. (1996). The imprint of somatic hypermutation on the repertoire of human germline V genes. *J. Mol. Biol.*, **256**, 813-817.

Tonegawa, S. (1983). Somatic generation of antibody diversity. *Nature*, **302**, 575-581.

Triglia, T., Peterson, M. G. and Kemp, D. J. (1988). A procedure for *in vitro* amplification of DNA segments that lie outside the boundaries of known sequences. *Nucleic Acids Res.*, **16**, 8186.

Tsumoto, K., Ogasahara, K., Ueda, Y., Watanbe, K., Yutani, K. and Kumagai, I. (1995). The role of Tyr residues in the contact region of antilysozyme monoclonal-antibody HyHel10 for antigen-binding. *J. Biol. Chem.*, **270**, 18551-18557.

Tsurushita, N., Fu, H. and Warren, C. (1996). Phage display vectors for in vivo recombination of heavy and light chain genes to make large combinatorial libraries. *Gene*, **172**, 59-63.

Tumas-Brundage, K. and Manser, T. (1997). The transcriptional promoter regulates hypermutation of the antibody heavy chain locus. *J. Exp. Med.* **185**, 239-250.

Ueda, T., Iwashita, H., Hashimoto, Y. and Imoto, T. (1996). Stabilization of lysozyme by introducing N-glycosylation signal sequence. *J. Biochem.*, **119**, 157-161.

Unanue, E. R. (1984). Antigen-presenting function of the macrophage. *Ann. Rev. Immunol.*, **2**, 395-428.

Valenzuela, P., Medina, A., Rutter, W. J., Ammerer, G. and Hall, B. D. (1982). Synthesis and assembly of hepatitis B virus surface antigen particles in yeast. *Nature*, **298**, 347-350.

van de Winkel, J. G. J., Bast, B. and de Groot, G. C. (1997). Immunotherapeutic potential of bispecific antibodies. *Immunol. Today,* **18**, 562-564.

van der Vaart, J. M., Te Biesebeke, R., Chapman, J. W., Toschka, H. Y., Klis, F. M. and Verrips, C. T. (1997). Comparsion of cell wall proteins of *Saccharomyces cerevisiae* as anchors for cell surface expression of heterologous proteins, **63**, 615-620.

van Deursen, J., Fornerod, M., van Rees, B. and Grosveld, G. (1995). Cre-mediated site-specific translocation between nonhomologous mouse chromosomes. *Proc. Natl. Acad. Sci. USA.*, **92**, 7376-7380.

Vaughan, T. J., Williams, A. D., Pritchard, K., Osboum, J. K., Pope, A. R., Earnshaw, J. C., McCafferty, J., Hodits, R. A., Wilton, J. and Johnson, K. S. (1996). Human antibodies with sub-nanomolar affinities isolated form a large non-immunized phage-display library. *Nature Biotech.*, **14**, 309-314.

Verhoeyen, M., Milstein, C. and Winter, G. (1988). Reshaping human antibodies: Grafting an anti-lysozyme activity. *Science*, **239**, 1534-1536.

Vonborstel, R. C., Ord, R. W., Stewart, S. P., Ritzel, R.G., Lee, G.S.F., Hennig, U.G.G. and Savage, E.A. (1993). The mutator mut7-1 of *Saccharomyces cerevisiae. Mutation Res.*, **289**,97-106.

Wach, A. (1996). PCR-synthesis of marker cassettes with long flanking homology regions for gene disruptions in *Saccharomyces cerevisiae. Yeast*, **12**, 259-265.

Wach, A., Brachat, A. and Philippsen, P. (1996). *Guidelines for Eurofan B0 Program: ORF deletants, plasmid tools, basic functional analyses.* Institute for Applied Microbiology, University of Basel, Switerland.

Wach, A., Brachat, A., Pöhlmann, R. and Philippsen, P. (1994). New heterologous modules for classical or PCR-based gene disruptions in *Saccharomyces cerevisiae. Yeast*, **10**, 1793-1808.

Ward, E. S. (1992). Antibody engineering: the use of *Escherichia coli* as an expression host, *FASEB*, **6**, 2422-2427.

Ward, E. S, Güssow, D., Griffiths, A. D., Jones, P. T. and Winter, G. (1989). Binding activities of a repertoire of single immunoglobulin variable domains secreted from *Escherichia coli. Nature*, **341**, 544-546.

Waterhouse, P., Griffiths, A. D., Johnson, K. S. and Winter, G. (1993). Combinatorial infection and *in vivo* recombination: a strategy for making large phage antibody repertories. *Nucleic Acids Res.*, **21**, 2265-2266.

Watzele, M., Klis, F. and Tanner, W. (1988). Purification and characterization of the inducible a-agglutinin of *Saccharomyces cerevisiae. EMBO*, **7**, 1483-1488.

Winston, F., Dollard, C. and Ricupero-Hovasse, S. L. (1995). Construction of a set of convenient *Saccahromyces cerevisiae* strains that are isogenic to S288C. *Yeast*, **11**, 53-55.

Winter, G. (1998). Synthetic human antibodies and a strategy for protein engineering. *FEBS letts.*, **430**, 92-94.

Winter, G., Griffiths, A. D., Hawkins, R. E., and Hoogenboom, H. R. (1994). Making antibodies by phage display technology. *Ann. Rev. Immunol.*, **12**, 433-455.

Winter, G. and Milstein, C. (1991). Man-made antibodies. *Nature,* **349**, 293-299.

Wojciechowicz, D. and Lipke, P. N. (1989). α-agglutinin expression in *Saccharomyces cerevisiae*. *Biochem. Biophys. Res. Commun.*,**161**, 45-51.

Wojciechowicz, D., Lu, C.-F., Kurjan, J. and Lipke, P. N. (1993). Cell surface anchorage and ligand-binding domains of the *Saccharomyces cerevisiae* cell adhesion protein α-agglutinin, a member of the immunoglobulin superfamily. *Mol. Cell. Biol.*,**13**, 2554-2563.

Wood, C.R., Boss, M. A., Kenten, J. H., Calvert, J. E., Roberts, N. A. and Emtage, J. S. (1985). The synthesis and *in vivo* assembly of functional antibodies in yeast. *Nature,* **314**, 446-449.

Wright, A. and Morrison, S. (1997). Effect of glycosylation on antibody function: implications for genetic engineering. *Tibtech.,* **15**, 26-32.

Wrighton, N. C., Farrell, F. X., Chang, R., Kashyap, A. K., Barbone, F. P., Mulcahy, L. S., Johnson, D. L., Barrett, R. W., Jolliffe, L. K. and Dower, W. J. (1996). Small peptides as potent mimetics of the protein hormone erythroprotein. *Science,* **273**, 458-463.

Yang. W. -P, Green, K., Pinz-Sweeney, S., Briones, A. T., Burton, D. R. and Barbas, C-F. (1995). CDR walking mutagenesis for the affinity matuaration of a potent human anti-HIV-1 antibody into the picomolar range. *J. Mol. Biol.,* **254**, 392-403.

Yang, Y. Y., Johnson, A. L., Johnston, L. H., Siede, W., Friedberg, E. C., Ramachandran, K. and Krunz, B. A. (1996). A mutation in a *Saccharomyces cerevisiae* gene (RAD3) required for nucleotide excision repair and transcription increases the efficiency of mismatch correction. *Genetics,* **144**, 459-466.

Yélamos, J., Klix, N., Goyenechea, B., Lozono, F. Chui, Y.L., González-Fernández, A., Pannell, R., Neuberger, M. S. and Milstein, C. (1995). Targeting of non-Ig sequences in place of the V segments by somatic hypermutation. *Nature*, **376**, 225-229.

Yelton, D. E., Rosok, M. J., Cruz, G., Cosand, W. L., Bajorath, J., Hellström, I., Hellström, K. E., Huse, W. D. and Glaser, S. M. (1995). Affinity maturation of the BR96 anti-carcinoma antibody by codon-based mutagenesis. *J. Immunol.*, **155**, 1994-2004.